

Disease insights from stem cells

Confronted by religious and political opposition, proponents of embryonic-stem-cell research need to draw attention not only to new therapies but also to the understanding of diseases that they could soon yield.

At a symposium on stem cells, regenerative medicine and cancer held last week at Stanford University in California, leading stem-cell researchers argued that the main effect of the technology on biomedical research will be felt not in cell-replacement therapies, but in applications that have until now received little attention: the use of embryonic stem (ES) cells in uncovering the mechanisms of genetic diseases, and in generating sources of normal and impaired tissues for use in drug discovery.

Human ES cells are derived from five-day-old human embryos, and their ability to grow and expand in an unlimited manner, and to form an array of specialized mature cell types upon command, has made them a valuable, albeit controversial, tool for biomedical research. A beautiful example of the power and limitations of ES cells was presented at the meeting by Hynek Wichterle, who works with Thomas Jessell at Columbia University in New York. He showed that mouse ES cells can be induced systematically to form motor neurons, the cell type that withers and dies in patients suffering from the fatal and incurable disease amyotrophic lateral sclerosis (ALS).

When he transplanted the motor neurons into mice, the cells survived and grew but failed to establish the appropriate connections within the mouse's nervous system. Nonetheless, the plates loaded with this precious cell type provide, for the first time, a plentiful and renewable source of motor neurons for identifying therapeutic molecules that promote motor-neuron growth, and prevent their death. Other scientists have developed protocols for inducing mouse and human ES cells to form a variety of tissues.

But even more powerful than generating healthy tissues from the existing human ES-cell lines would be the ability to generate new cell lines that could produce unlimited quantities of diseased tissue where the disease has a genetic basis, such as motor neurons from ALS patients, or cardiac muscle cells from heart patients. Scientists could then use the cell lines to study the cellular processes that go awry in the diseased cells, and design high-throughput screens to identify molecules that halt disease progression. It is feasible to generate ES-cell lines from mouse models of human disease, but often the underlying causes may be different from those of the human diseases they are supposed to represent. And many human diseases are so complex or poorly understood that mouse models do not exist for them.

Growth area

It has not yet been possible to generate ES-cell lines from adult humans — the existing lines have all been derived from embryos left over from *in vitro* fertilization (IVF). But although this appears to be technically difficult, it will no doubt be worth the effort. For example, Irving Weissman, who heads the newly launched Institute for Cancer/Stem Cell Biology and Medicine at Stanford University, has made this one of the goals of the privately funded institute, and plans to generate new ES-cell lines from patients with cancer, diabetes, cardiovascular disease, autoimmune disease, allergies and neurodegenerative diseases such as Parkinson's and ALS.

Weissman argues that the existing human ES-cell lines, which are all derived from IVF embryos, do not harbour the mutations that cause these diseases and so cannot be used for studying human disease in

the manner that he proposes. Moreover, the existing cell lines do not reflect the genetic diversity of the overall population, but rather that of the mainly white, infertile clientele of the clinics. These lines could not help scientists to understand, for example, why there is currently an epidemic of type II diabetes in the Native American population.

With current technologies, generating new ES-cell lines from patients would require somatic-cell nuclear transfer (SCNT). This procedure involves extracting the nucleus from a patient's mature cells (skin cells, for example). These cells will contain virtually identical genetic information to that contained in the patient's diseased cells. The nuclei would then be transferred to an enucleated egg (oocyte), which would be stimulated to divide in a Petri dish to develop into an embryo. After about five days, the ES cells would be extracted from the blastocyst (which destroys it), cultured *in vitro* and induced to form the diseased tissue.

Obstacles

Although using ES-cell technology to understand the mechanisms of human disease could be spectacularly illuminating, there are practical and political hurdles in bringing this idea to fruition. The first problem is that deriving new ES cells from human embryos, even spare IVF embryos that are destined for destruction, is not allowed in the United States with public funding under current federal guidelines — although it is permitted if funds are private. It is also forbidden in several European countries.

Second, even if the technical problems of performing SCNT with human cells is overcome, there is opposition to SCNT because it involves the production of human embryos for research. Moreover, it has inappropriately been lumped into the general category of 'cloning', and so may be subject to regulation that is aimed mainly at preventing reproductive cloning. If the Brownback Bill pending in the Senate is passed, US researchers who perform this procedure could be subject to 10 years' imprisonment and a \$1-million fine. And if amendments to a safety bill last week by European parliamentarians are ratified by national governments, such procedures would be banned in Europe, along with all other techniques for producing embryos for research. In contrast, Singapore allows the use of SCNT to produce human embryos, which can then be kept for 14 days to allow ES cells to be extracted.

There are practical difficulties too. Human oocytes for SCNT are in short supply because of the invasive nature of the procedure required to harvest them, so scientists will need to find a way to collect eggs from ovaries that have been donated for research (each of which contains about 200,000 oocytes), learn to culture human oocytes from progenitor cells, or develop ways to generate human ES-cell lines without using oocytes.

Although there are numerous challenges that have yet to be overcome, a 'tipping point' will eventually be reached where technical problems will become solvable. This will result in a sea change in the field, leading to rapid progress. The implications of ES-cell technology for understanding and curing human disease are enormous. These benefits need to be fully appreciated and communicated in the midst of political and ethical debates. ■





Taking the strain
Database set up
to combat drug-
resistant bacteria
p791



Deep freeze
Tonnes of Atlantic
cod killed by icy
Canadian waters
p792



Truth test
US refuses to pull the
plug on lie detectors
for lab scientists
p793



Summing up
Mathematicians
excited by potential
of Poincaré proof
p794

Historians pool resources to halt trade in Iraq's stolen treasures

Hannah Hoag, Washington

Archaeologists are rallying in an international effort to create a catalogue of the objects stolen from the National Archaeological Museum in Baghdad earlier this month. Images and descriptions of the artefacts will be entered into a database to help law-enforcement agents crack down on illegal trafficking.

Within days of US forces entering Baghdad, looters ransacked the museum in what many archaeologists suspect was an organized raid. The museum held some 180,000 artefacts including cuneiform tablets, statues, pottery and jewellery.

McGuire Gibson, an archaeologist at the University of Chicago's Oriental Institute, is spearheading production of the catalogue. Graduate students at the institute are scanning images and extracting detailed information about each object from archived expedition records, manuscripts, books and catalogues. The effort is being supported by universities and institutions in the United States, Britain, Germany, France and Japan.



A shattered statue lies among the debris in Baghdad's museum following the recent looting spree.

"Everything that was excavated in Iraq since the 1970s was in the museum," says Elizabeth Stone, an archaeologist at the State University of New York in Stony Brook.

Gibson was one of 30 international experts who last week convened at the Paris headquarters of UNESCO, the UN Educational, Scientific and Cultural Organization,

to coordinate efforts to compile information on the lost collection. UNESCO director-general Koïchiro Matsuura told the meeting that he would seek a UN resolution to impose an embargo on the acquisition of all Iraqi cultural objects, and demand the return of those already exported.

Many archaeologists are pushing for a more extensive embargo that would temporarily ban the trade of all Middle Eastern antiquities, and for amnesty and rewards for Iraqis who return stolen goods. The latter proposal is "controversial, but we have to be pragmatic", says archaeologist Tony Wilkinson of the University of Chicago.

The fast reaction of the archaeologists is an attempt to salvage a situation they originally sought to prevent. Gibson visited the Pentagon in January, when he stressed the importance of the Baghdad museum. "I told them that it was likely to be looted and that the army would have to move in fast if the guards ran," he says. He says that Pentagon officials assured him that the museum would be protected.

Last week, three members of President Bush's Cultural Property Advisory Committee, which advises the White House on measures to prevent the illicit trade in cultural property, resigned in response to the looting. In his letter of resignation, chairman Martin Sullivan cited the failure of the US military forces "to plan for and to try to prevent indiscriminate looting and destruction" of the museum.

US debates future of Iraqi scientists

Erika Check, Washington

Ten years after the South African bioweapons programme was officially dismantled, some of its expertise and materials are still at large, prompting concern about similar programmes in other nations, such as Iraq.

The Washington Post reported on 20 April that Daan Goosen, a founding director of the lab that did research for South Africa's bioweapons programme, offered the FBI a vial of genetically modified *Escherichia coli* bacteria. According to the newspaper, he promised to turn over a cache of other weapons in exchange for money and a US passport, but was refused.

The fact that scientists from the programme are still seeking to exploit their knowledge alarms non-proliferation experts who are now looking at the situation in Iraq.

Jonathan Tucker, a senior fellow at the US Institute of Peace, a government-funded think-tank in Washington DC, says that the Bush administration should establish programmes in Iraq similar to those that the government has set up to employ former weapons scientists in Russia and the Ukraine. Earlier this month, Congress rejected a proposal to spend \$50 million on such 'cooperative threat reduction' programmes outside the former Soviet Union.

Others experts says that such programmes won't be necessary in Iraq, whose biological-weapons programme was far less extensive than that in Russia. "The concern will be, can we encourage these scientists to stay in Iraq and help build a biotechnology industry which is good for their country?" says David Franz, a former United Nations weapons inspector.

India pledges deluge of funds to water plan

K. S. Jayaraman, New Delhi

A court judgement and last year's dry monsoon season have injected fresh momentum into a massive water project that would integrate most of India's major waterways at a cost of more than US\$100 billion.

The project would link 37 rivers through a maze of 30 canals totalling 12,500 kilometres in length. It would involve the construction of 12 reservoirs, displacing an estimated 450,000 people and flooding 80,000 hectares of forest.

India's Supreme Court ruled late last year that the government should meet its constitutional requirement to supply clean water to each Indian citizen by 2016. After the ruling, Prime Minister Atal Behari Vajpayee pledged that the project should be undertaken on a "war footing" and completed by that year.

Now Indian scientists and engineers are preparing detailed plans to implement the project, the broad outline of which has languished at the planning stage for more than 20 years. But the scheme still faces several political obstacles at home and abroad, and some environmentalists are already attacking it.

"The idea is to transfer surplus water from Himalayan and other rivers to regions where it is scarce for drinking and agriculture," says Suresh Prabhu, an engineer and former energy minister who heads a task force set up after the court judgement to implement the project.

But the project would require most of India's states — as well as the parliament in New Delhi — to ratify a constitutional amendment transferring responsibility for water supply to the central government. And moves to take more water from two huge rivers involved in the scheme — the Ganges and the Brahmaputra — will raise problems with India's neighbour Bangladesh, which relies heavily on water from each of them (see *Nature* 422, 254–256; 2003).

Of the 1.9 trillion cubic metres of rainwater that pass through India's rivers in a typical year, about two-thirds flow into the sea. This water could, at least in theory, be used to bolster the subcontinent's water supply. Huge transfers of water between different river basins "may be the best way to efficiently utilize the national resource", argues Pavagada Venkata Indiresan, a project supporter and former president of the Indian National Academy of Engineering.

The proposed network would deliver up to 173 billion cubic metres of water — equivalent to a quarter of the Brahmaputra's flow — each year, enough to irrigate 35 million hectares and generate up to 34 gigawatts of electricity, according to Prabhu.

The project has found broad support, both among India's political parties and



Pipe dream: India plans to harness rivers such as the Ganges to improve access to clean water.

within its scientific community. "It can result in a win-win situation for all the states," enthuses Monkombu Swaminathan, the agricultural scientist who is regarded as the father of India's green revolution.

The government has asked the Indian Institutes of Technology (IIT) to create simulations of the proposed waterway system, and has requested data from remote-sensing satellites operated by the Indian Space Research Organization. These data will be

used to help to align the canals, locate suitable sites for dams, and avoid areas of likely ecological damage.

But critics, such as A. Vaidyanathan of the Madras Institute of Development Studies in Chennai, argue that the project will damage natural river ecosystems and that the government would be better off improving traditional techniques of rainwater collection and increasing the efficiency of existing irrigation systems.

"Linking the rivers is a mirage," warns Ramaswami Iyer, a former secretary of water resources, who says that the costs and political objections surrounding the project are insurmountable.

Prabhu says that the IIT will "factor the ecological fears into their simulation models" in planning the project. He also says that the project's latest estimated cost of \$112 billion can be raised through water charges, international loans and private-sector contributions. "We need 1% of the gross domestic product (GDP) every year, and once implemented, the project will boost our GDP by 5%," he predicts. ■

Biotech firms pin hopes on defence

Helen Pearson, New York

Biotechnology companies — battling for survival after three years of waning investment — are turning to the US government's war on bioterrorism as a source of much-needed revenue.

The companies are eyeing initiatives such as Project BioShield, which will provide an estimated \$6 billion over ten years for the development of drugs and vaccines against agents such as anthrax and smallpox.

"It's an opportunity created by fear," says Charles Cantor, a biodefence expert and chief scientific officer at Sequenom, a genomics company based in San Diego, California.

VaxGen of Brisbane, California, suffered a blow to its long-standing objective of developing an AIDS vaccine in February, when it announced disappointing clinical-trial results for such a vaccine (see *Nature* 421, 877; 2003). But last October, the company was awarded \$13.6 million by the National Institute of Allergy and Infectious Diseases for trials of a vaccine against anthrax. "You have to shift with the times," says Lance Ignon, a company spokesman.

SIGA Technologies of Corvallis, Oregon, is hoping that the government will support the development of its prototype vaccines, which use harmless bacteria that are genetically engineered to produce anthrax, smallpox or plague molecules to stimulate

an immune response. The company hopes to take advantage of a change made in June last year to the Food and Drug Administration's regulations. Under the new rules, vaccines against bioweapons can be approved on the basis of animal experiments alone.

Apart from generating revenue, bioterrorism projects are helping to shore up investors' confidence in some biotechnology stocks. Last month, shares in Human Genome Sciences of Rockville, Maryland, went up by about a quarter when the company announced the success of an antibody-based drug for anthrax in trials using rabbits and monkeys.

Industry analysts warn that biodefence research is unlikely to fuel a long-term revival in biotechnology investment. "I don't think it'll pull them out of a hole," says analyst Stephen Burrill of Burrill and Co., a San Francisco investment bank that specializes in biotechnology.

Part of the problem is that there is no sustainable market for biodefence drugs and vaccines after the government has stockpiled them, says Jeff Bird of venture-capital firm Sutter Hill Ventures in Palo Alto, California. And with many companies plumping for drugs and vaccines against the same diseases, they will have to compete with each other to supply the stocks. ■

Biologists deploy database to quash drug-resistant bacteria

Tom Clarke

European microbiologists have assembled a database to help combat the rising tide of antibiotic-resistant bacteria.

A multinational effort involving 12 research centres has collected information on the known European strains of methicillin-resistant *Staphylococcus aureus* (MRSA), and has standardized the technique for identifying the bacteria (S. Murchan *et al.* *J. Clin. Microbiol.* **41**, 1574–1585; 2003).

“We’ve mapped the existing distribution of strains and followed the current spread of others,” says Barry Cookson, a microbiologist at London’s Central Public Health Laboratory (CPHL) who led the initiative. The results confirm reports that some strains of MRSA have managed to jump between hospitals, both within Europe and farther afield.

Although normal strains of *S. aureus* are common and relatively harmless, MRSA has become a significant problem in hospitals, where it tends to prosper because of the pervasive use of antibiotics. The bacteria can infect the lungs, bones and bloodstream of patients, and are responsible for thousands of deaths worldwide each year.

In the past year, MRSA strains have also begun to appear that are resistant to vancomycin, the antibiotic used by most hospitals as their last resort, as well as to the completely new antibiotic linezolid. In addition, outbreaks of MRSA infections that are not associated with hospitals are becoming more common.

Many hospitals have experience of treating their native strains of MRSA using cocktails of various antibiotics. But the database should allow them to identify any new strain early, which could prevent the invaders from gaining a foothold. “People will be able to identify and communicate what could be a common problem,” says Alex van Belkum, a microbiologist at Erasmus University Medical Center in Rotterdam, the Netherlands, one of the paper’s authors.

The identification method recommended by the European team is pulsed-field gel electrophoresis (PFGE). This identifies strains based on how far fragments of their DNA move through a gel under the influence of electrical pulses. But minute variations in the gel’s preparation and the electric field can give variable readings, the researchers say.

PFGE is so sensitive that results can vary “depending on whether it’s raining or not”, jokes microbiologist Tyrone Pitt at the CPHL. To overcome this, the team has specified the amount of gel to be used, the concentration of DNA in the sample, the temperature, and the timing of the electrical pulses.



Caught: drug-resistant strains of *Staphylococcus aureus* can now be identified more readily.

Gel images from any lab can now be matched with strains held on the database and reliably identified, the researchers say.

The European team is now pushing to get the standard PFGE protocols adopted by the International Union of Microbiological Societies. Global standardization and a central database could be valuable in preventing the spread of the bacteria. Already MRSA strains first isolated in Europe have reached Australia and New Zealand, where hospitals have begun screening doctors visiting from parts of Europe for MRSA in case they have brought the bacteria with them.

The US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, is also assembling an MRSA database. “We’ll have it up and running by the end of the year,” says CDC microbiologist Fred Tenover. Initial work between the CDC and researchers in the Netherlands suggests that the two PFGE protocols will be roughly comparable. “It’s a bit ad hoc but we are certainly able to share information,” says Tenover.

Both the European consortium and the CDC are also developing PFGE protocols for antibiotic-resistant *Pseudomonas aeruginosa* and *Enterococcus* bacteria.

In the long term, PFGE tests are likely to be replaced by definitive genome-sequence data for each strain of bacteria. But for now, most hospitals and public-health labs lack the means to produce such sequence data for strains that they want to check.

National Academies launch grants for interdisciplinary work

Erika Check, Washington

The US National Academies are set to start awarding grants of their own for the first time, under a \$40-million project funded by the California-based W. M. Keck Foundation.

The National Academies Keck Futures Initiative will focus on interdisciplinary research, says Ken Fulton, who will head it. As part of the initiative, the academies will host small meetings to bring together researchers from various disciplines to discuss particular themes.

Researchers who attend the meetings may subsequently apply for grants worth \$200,000 to support interdisciplinary projects. The programme is funded for 15 years, and four seed grants are expected to be given during its first four years.

“We’re hoping to foster innovative, imaginative cross-disciplinary work that might not have occurred without the interaction of the investigators at these conferences,” Fulton says.

Conferences will be held twice a year, beginning this November with a meeting on signalling in neuroscience, cell biology and engineering. Participants will be asked to discuss how signalling networks convey information in physical and biological systems, and how these systems can inform each other. Future topics will include nanotechnology and genomics.

Although the money will only support a few researchers, a spokeswoman for the Keck Foundation says that the initiative hopes to reach a broader audience through the conferences. The project is also intended to affect the research direction of funding agencies such as the National Science Foundation and the National Institutes of Health, which have often been criticized as slow to support multidisciplinary work. Representatives from these government agencies will be invited to the meetings, she adds.

“We’re trying to get scientists from across disciplines to sit down and ask new questions, rather than having the biologists ask the mathematicians to help them solve a piece of the problem in some research they’re conducting,” says a spokesman for the Keck Foundation. The futures initiative will fund a National Academies consensus study on how best to foster interdisciplinary research.

The academies will also rename their new building in Washington DC, which opened last year. On 13 May, the building will be christened the Keck Center of the National Academies.

University faces up to wartime use of slaves

Alison Abbott, Munich

After two years of bickering, the University of Göttingen has given the go-ahead for a project to establish the fate of slave labourers employed and treated at its medical faculty under the Third Reich.

The university has also agreed to publish a summary of the findings of an initial investigation led by historian Andreas Frewer, who delivered his report to the faculty's board last June. And it is calling for donations from staff and students to an existing compensation fund for the slave labourers.

Frewer has long been at loggerheads with the board of the medical school, claiming that its members have hindered his work. He was appointed by the university's Department of Medical Ethics and History of Medicine at the beginning of 2000, with a brief to investigate whether or not the university clinics had exploited slave labour, as many

industrial employers did during the Second World War.

In his initial study, Frewer and his colleagues located files from the university's surgical and gynaecological wards, and identified 125 slave labourers who had worked within the faculty between 1940 and 1945. They also found details of some 600 slave labourers who were treated in the wards.

The records were used as evidence by some former slave labourers who applied to a federal compensation fund that was established in Germany in 1999.

But Frewer was denied access to neurological and psychiatric files on the patients, which were kept by a private company. Moreover, he claims, the faculty board obstructed his path, slowing down the publication of his results, limiting the amount of assistance that he received, and failing to respond to requests for meetings. The board



Slavery in Würzburg, 1945: university historians are now unearthing its true extent in Germany.

also delayed and watered down press releases about the results, he alleges, and failed to respond to the project's final report.

Frewer argues that if the medical faculty had been more dedicated to the project, many more of the slave labourers — whose numbers are now rapidly dwindling — would have been able to claim compensation from a state fund, which is now no longer accepting claims. He also criticizes the faculty for not contributing to the federal fund.

The project's difficulties have also been noted by outsiders. Ernst Böhme, head of the Göttingen City Archive, which has also sought and supplied documents to support compensation claims by former slave labourers, says that the historical research at the medical faculty seems to have been proceeding too slowly.

But Manfred Droese, dean of the faculty, denies that he or his board sought to hinder the project or to block discussions of its results. "Maybe we caused a bit of delay, but I started the project because we wanted to face the faculty's history," he argues. He adds, however, that the project should have concentrated on documenting the employment of slave labourers, rather than devoting effort to assisting compensation claims.

The neurological and psychiatric files contain sensitive information about people who could still be alive, Droese says, and their examination had to be approved by experts in data-protection law. But he says that the files will now be made available for a new research project.

According to Claudia Wiesemann, director of the university's ethics department, the new project will have a broader scope in looking for what happened generally to slave labourers in psychiatric and neurological clinics, and will attempt to determine how many were sterilized or killed.

Droese describes Frewer as 'self-righteous', but Frewer's supporters say that it is only his perseverance that has forced the faculty board to move the research forward.

HULTON ARCHIVE

Atlantic cod meet icy death

Hannah Hoag, Washington

More than 700 tonnes of Atlantic cod have frozen to death in chilly waters off eastern Newfoundland.

The dead fish first began to surface in the waters of Smith Sound, Trinity Bay, on 3 April, but scientists are struggling to explain what induced the cod to swim into the lethal waters.

Such mass culls are very rare, although one in the northeastern United States killed an estimated 10 million to 100 million 'warm water' tilefish fish in 1882 (R. Marsh *et al. Fish Oceanogr.* 8, 39–49; 1999).

At this time of year, cod are usually found in warmer waters near the bay. Local theories on why they entered the sound include a bid to escape from seals and a fateful pursuit of herring. But none of these ideas passes muster, says George Lilly, a biologist with Canada's Department of Fisheries and Oceans (DFO), who is investigating the deaths.

A team of biologists and oceanographers from local universities and the DFO has been monitoring water temperature, salinity and oxygen saturation in the sound since last week. They have also collected dead cod for analysis and confirmed that they were perfectly healthy when frozen. They have used acoustic and video monitoring equipment to track down survivors.

In early April the temperature of the water column in Smith Sound fell to -1.7°C , says Eugene Colbourne, an oceanographer with the DFO. Historical temperature profiles from the region indicate that such temperatures are very unusual for the sound.

Cod produce antifreeze proteins in their blood to safeguard them from very cold water, but the proteins take two months to build up to maximum levels and the dead fish had very little of them in their blood, say investigators.

Some environmentalists are describing the mass kill as an environmental disaster, as it destroyed one of the last remnants of the region's few cod stocks.

Canada's fisheries minister, Robert Thibault, is expected to make an announcement shortly on Newfoundland's fisheries, and may ban cod fishing. Cod used to make up 70–80% of the total stock in the province, says Jeffrey Hutchings, an evolutionary biologist at Dalhousie University in Halifax, Nova Scotia, but 99.9% of the stock has been lost. North Atlantic cod is classed as being of 'special concern' by the Committee on the Status of Endangered Wildlife in Canada.



Chilling effect: a mass kill has destroyed hundreds of tonnes of endangered Canadian cod.

Lie probes to continue at US weapons labs

Geoff Brumfiel, Washington

The US Department of Energy (DOE) has said that it plans to continue routine polygraph testing of scientists who have access to classified information in its nuclear-weapons laboratories.

The decision has angered many of the researchers, who say that it ignores a National Academy of Sciences study published last October, which found that polygraph testing has "extremely serious limitations" in discriminating between honest employees and likely spies.

Announcing the plan, energy secretary Spencer Abraham cited both the war in Iraq and the war on terrorism as reasons for continuing with the polygraphs. "As the steward of the nation's nuclear-weapons stockpile, the department has an obligation to use the best tools available to protect its most sensitive information," he said in a statement on 14 April.

But critics say that the National Academy study found polygraphs to be an unscientific tool that is unlikely to detect spies. "It's lore, it's theology, and unfortunately, the secretary of energy has bought it," says Alan Zelicoff, a bioweapons specialist at the Center for National Security and Arms Control at Sandia National Laboratory in Albuquerque, New Mexico. "I can't tell you the number of e-mails I've got today saying, 'This represents a new low'," he adds.

The DOE introduced routine polygraphs for thousands of staff at its nuclear-weapons

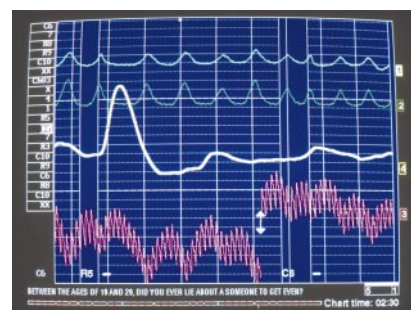
labs three years ago as a response to the case of Wen Ho Lee, an engineer at the Los Alamos National Laboratory in New Mexico, who was charged with passing nuclear secrets to a foreign power. Lee was not convicted, but pleaded guilty to lesser charges.

Counterintelligence officers at the department argued that polygraphs would help them unearth potential spies, but many scientists pointed to a lack of data that the tests actually work.

Their scepticism led Congress to commission the National Academy study, which found that polygraphs could sometimes detect deceitful responses to directed questions, but lacked the precision to identify lies in response to the sort of open-ended questions asked in security-clearance interviews. The study concluded that testing left officials with an "unacceptable choice" between falsely judging too many loyal employees or leaving security threats undetected.

Despite these conclusions, the DOE plans to continue the tests. "It would not be appropriate to downgrade one of the most effective tools we have," says DOE spokeswoman Jeanne Lopatto. She adds that even if the study questioned the polygraph's utility, it still suggested that the tool can be effective if used as a trigger for a detailed follow-up investigation.

The study's authors disagree. "The report is clearly taken out of context and distorted," says Stephen Fienberg, a professor of computer science and statistics at Carnegie Mellon University in Pittsburgh, Pennsylvania, who



Not the end of the line: the US energy department is defying critics of its routine polygraph screens.

chaired the National Academy study. "There is virtually no science behind the screening process on which they rely," he adds.

Politicians representing the areas around the weapons labs are also unconvinced. In a statement, Congresswoman Ellen Tauscher (Democrat, California) said that she was "shocked" by the decision and called for hearings. Senator Pete Domenici (Republican, New Mexico) and Jeff Bingaman (Democrat, New Mexico) also voiced concern at the plan, and may hold a hearing after a period of public comment on it ends on 13 June.

"I'm disappointed by the decision," says Jeff Colvin, a physicist at Lawrence Livermore National Laboratory in California and an officer for the lab's union for scientists and engineers, which opposes the screening. "It's clear that the DOE will not be swayed by logical arguments or scientific facts."

Agent Orange health investigation stuck at square one

Declan Butler, Paris

A full year after the United States and Vietnam agreed the first bilateral research programme to investigate the health and environmental damage caused by the defoliant Agent Orange, the initiative remains stuck on the starting blocks.

The United States and Vietnam signed

a memo formally establishing the joint programme in March last year, following a four-day conference in Hanoi that brought together epidemiologists, toxicologists and environmental scientists from 13 countries (see *Nature* 416, 252; 2002).

But so far the two nations have not even set up the joint advisory committee needed to run the initiative, approve areas for support, and thrash out the ground rules for the project. These would include ethical guidelines, publication and authorship policies, and allocation of resources for training and equipment in Vietnam. Researchers fear that any lapse in the initiative could stop them following up on new, more specific findings about where Agent Orange was sprayed by US forces during the Vietnam War (see *Nature* 422, 649; 2003).

US officials associated with the programme accuse Vietnamese officials of failing to get the plan moving. "The United

States proposed its members of the joint committee last June and invited Vietnam to do the same, but we have not yet had a response," says Anne Sassaman, a director at the US National Institute of Environmental Health Sciences and a signatory of the memo along with Nguyen Ngoc Sinh, director of Vietnam's National Environmental Agency.

Sassaman says that a Vietnamese delegation visited the United States last September, and she herself went to Vietnam the following month to get the project on track. But she says that "since then we have had no communication with Vietnamese officials". Last month, the US ambassador in Hanoi, Raymond Burghardt, wrote to deputy prime minister Nguyen Tan Dung, asking him to move the initiative forward, according to Sassaman.

Vietnamese researchers and officials who are involved in the project, including Sinh, did not respond to requests for information on the plan's current status in Vietnam.



NASA feels the heat as glacier pictures cause unrest in Peru

Lima NASA has come under fire for causing unnecessary alarm with a recent press release warning of an “ominous crack” that has appeared in an Andean glacier.

An infrared imaging device on Terra, a NASA Earth-observation satellite, identified the crack, which was detailed in a press release on 11 April. The release warned that if the glacier broke off and fell into the lake that it feeds, it could cause a flood that would reach the 60,000-strong Peruvian town of Huaraz in less than 15 minutes. Many inhabitants of Huaraz became alarmed after the release was covered in local newspapers, which referred to earlier, similar events in which thousands died in the region.

NASA says that the press release was intended to promote the value of remote sensing in monitoring potentially dangerous geological conditions, but Earth-observation experts have questioned whether it was responsible to do this, as the glacier posed no immediate threat. José Achache, director of the European Space Agency's Earth Observation Programmes, says it is “not ethical for an agency to refer to threats to people's lives as a means of demonstrating the relevance of Earth-observation satellites”.



Warnings over satellite pictures of a glacier have unsettled the residents of Huaraz (bottom left).

Stench of fraud in German odour lab as paper cries foul

Munich The atmosphere is not so fragrant in Germany's top lab for the physiology of smell, according to the *Süddeutsche Zeitung*.

Allegations of data manipulation have been brought against its head, Heinz Breer of the Hohenheim University near Stuttgart, concerning at least two publications from the past five years. The newspaper reports this week that in one paper (S. Schreiber, J. Fleischer, H. Breer and I. Boekhoff *J. Biol. Chem.* 275, 24115–24123; 2000), a photograph of a western blot is alleged to have been manipulated. It also claims that some results in another paper (J. Noé and H. Breer

Poincaré proof adds up to potential payday

Moscow Mathematicians around the world are poring over a proof of the Poincaré Conjecture, one of seven US\$1-million Millennium Prize Problems.

Grigori Perelman of the Steklov Mathematical Institute of the Russian Academy of Sciences in St Petersburg released two versions of the proof as preprints, one in November and the other last month. The conjecture concerns three-dimensional surfaces. Closed two-dimensional surfaces without holes can be transformed onto the surface of a sphere, and Henri Poincaré (pictured) suggested that similar surfaces with higher dimensions should also transform back to spheres.

In 2000 the Clay Mathematics Institute in Cambridge, Massachusetts, offered an award of \$1 million to anyone providing a proof of the



conjecture that could withstand academic scrutiny for two years. Six other problems were allocated an equal amount. Proofs have appeared frequently on preprint servers (see *Nature* 416, 670; 2002), but have never survived academic scrutiny. But Perelman's highly

technical paper is considered very promising, and many universities are planning seminars in which researchers will probe the proof for errors.

► www.arxiv.org/abs/math.DG/0303109

J. Neurochem. 71, 2286–2293; 1998) seem to have been altered. Breer told the paper that “something was not quite right” but denied any intention to deceive.

In 1998, the DFG, Germany's granting agency for universities, awarded Breer the country's most prestigious research prize, the Leibniz award. The agency is now investigating the affair.

Japanese lab turns the page on cDNA delivery

Tokyo Geneticist Yoshihide Hayashizaki is going back to basics in his bid to supply researchers with samples from his laboratory. From this summer, Hayashizaki plans to send out his collection of 60,000 mouse complementary DNAs (cDNAs) in the form of a hardcover book.

The cDNA samples, which can be used in gene-expression studies, are in great demand from researchers around the world (see *Nature* 420, 602–604; 2002). Hayashizaki, who works at RIKEN's Genomic Sciences Center in Yokohama, says the DNA book will be mailed by courier, cutting the time and costs incurred by the current supply system, which involves shipping a large box of refrigerated samples.

When researchers drop the cDNA of choice into a tube of water, the paper dissolves leaving the cDNA behind. The samples have been shown to last for about four months, although Hayashizaki says they should last for over a year. His group is patenting the printing method.

AIDS scientist rapped over controversial treatment test

Los Angeles A researcher at the University of California, Los Angeles (UCLA), broke federal human-protection rules by “indirectly” participating in tests of a controversial AIDS treatment known as malaria therapy, the university said on 16 April.

John Fahey became involved in the trial in 1997 when he supervised Xiao Peng Chen, who brought samples to UCLA from a Chinese malaria therapy trial. The widely discredited approach, which involves deliberately infecting AIDS patients with the parasite that causes malaria, is championed by Henry Heimlich, inventor of the famous manoeuvre to aid choking victims.

“Fahey and his laboratory participated only indirectly in the research by allowing the testing of Chen's samples,” the university said in a statement. But because UCLA's Medical Institutional Review Board had not approved the research, Fahey was found to be in violation of regulations on the use human subjects. The university has referred Fahey's case to the federal office that oversees human research protections.

Congo massacre probe enlists forensic team

London Forensic-science experts arrived this week in the Democratic Republic of the Congo as part of a 15-strong United Nations team charged with investigating the alleged massacre of 350 civilians.

The killings occurred on 3 April in the Ituri district, where a war within a war rages between local militias. Forensic investigation will be a first step towards identifying the perpetrators, according to the Office of the High Commissioner for Human Rights at the United Nations. An estimated 50,000 people have been killed in and around Ituri since 1999. The day before the massacre, the government, rebel groups and opposition parties agreed on a fragile deal to end the country's four-year civil war.

The human-rights office is also updating its list of forensic experts to feature a broader range of specialists, in the light of recent investigations of atrocities in Afghanistan. It currently draws on medical experts from crime labs and academic forensic anthropologists and archaeologists.



A recipe for revolution?

Sequencing the DNA of the world's leading food crop was the easy part. Now comes the tricky task of turning our new knowledge of the rice genome into agricultural and economic gains. David Cyranoski reports.

Last year was a good one for rice genomics. Draft genome sequences of the two agriculturally important subspecies of rice, called *indica* and *japonica*, were published in April^{1,2}. And in November, the International Rice Genome Sequencing Project (IRGSP) unveiled high-quality sequences of two of *japonica*'s 12 chromosomes^{3,4}.

In the wake of these achievements, expectations are high. "Rice DNA finding will transform how the world is fed," is how one British newspaper reported the publication of the two draft sequences. But what do we really know about the rice genome, and its potential agricultural benefits? Not much, admits Takuji Sasaki of the National Institute of Agrobiological Sciences in Tsukuba, Japan, who heads the IRGSP. "We are at the starting line for rice genomics, both basic and applied," he says.

From the standpoint of Sasaki and his IRGSP colleagues, the first stage of this race will involve identifying each rice gene and assigning functions to them. But whereas the sequencing was conducted as an entirely open, team effort, economic considerations may mean that rice functional genomics will become a rather less collaborative venture. Some governments that are investing in the field want to ensure that their own nationals have privileged access to any tools that they develop, and intellectual-property issues complicate the picture still further.

There is no clear finishing line, but the worth of the rice genome will eventually be judged in terms of economic and agronomic gains. This means that different researchers may end up running in different directions. Rich countries such as Japan, for instance, are interested in improving such traits as the taste and texture of the grain. In contrast, plant breeders at the International Rice Research Institute (IRRI) near Manila in the Philippines want to produce higher-yielding or nutritionally superior varieties that will be able to tolerate harsh environmental conditions, in order to improve the lot of impoverished farmers in the developing world.

Two cultures

What's more, achieving these divergent goals will require genome researchers to begin a meaningful dialogue with breeders in the field. Currently, the different scientific cultures and approaches of the two groups present a formidable obstacle. "These two communities have to get together, but it's like there's a bridge missing," Susan McCouch, who works on molecular approaches to plant breeding at Cornell University in Ithaca, New York, told the International Rice Genome Meeting 2003, held in Tsukuba in February. All in all, it seems that researchers in different camps need to figure out what they want from the rice genome, and how best to get it.

At least the genome researchers are clear on their first move: to work out the functions of all

Rice to the occasion: farmers in poor nations such as Nepal (above) and rich ones like Japan (inset) stand to gain different things from rice genomics.

of the rice genes — a total estimated at around 60,000 by gene-hunting computer programs. About half of these genes have been assigned vague functions on the basis of their sequences — researchers might surmise, for instance, that a gene encodes a member of a particular class of enzymes. But much work remains to be done. "The categories are almost meaningless," says Hirohiko Hirochika of Japan's National Institute of Agrobiological Sciences. So far, only about 100 rice genes have been ascribed a precise, verified function.

The availability of sequence information has already hastened the gene hunt, however. Masahiro Yano, also at the National Institute of Agrobiological Sciences, estimates that the rice genome project trimmed between one and three years off his hunt for a gene that controls flowering time⁵, by providing genetic markers that he could track through breeding experiments to pin down the location of a candidate gene.

Accelerating progress further will depend on new tools, including huge libraries of mutant plants created by randomly inserting tagged bits of DNA into the genome to disrupt their genes. Hirochika has already created 50,000 mutants using this method, whereas Gynheung An at the Pohang University of Science and Technology in South Korea has

made a library of 100,000 plants. In half of An's mutants, the tagged DNA contains a promoter sequence that can boost the activity of nearby genes. An's team has already created plants in which growth is stunted or flowers bloom late, as well as mutants with increased sensitivity to heavy metals or salts⁶.

In addition to randomly created mutants, rice geneticists would like to develop the ability to knock out specific genes at will. A new tool developed by Shigeru Iida and his colleagues at the National Institute for Basic Biology in Okayama, Japan, could provide the answer⁷. It exploits the occasional tendency of DNA to insert itself into the genome at points where the sequence matches its own — a phenomenon known as 'homologous recombination' — to disrupt the function of particular target genes.

DNA microarrays will also speed the analysis of gene function. Microarrays can carry tens of thousands of genes or predicted gene sequences, allowing quick screens of a plant's genetic make-up, or the activity of its genes, to be carried out. A chip made by the multinational agribiotech firm Syngenta has already been used to identify 269 genes that are expressed when a rice grain is being filled up with its store of carbohydrates, proteins and fatty acids⁸. Microarrays are especially useful because they can reveal the activity of networks of genes, which are extremely difficult to study by random mutation or targeted gene knockouts.

But access to such tools is not completely open. Syngenta's microarray is a proprietary technology, and Iida's gene-targeting technique is similarly being patented. Newly discovered genes may also be patented as they are discovered. "Intellectual property came into rice research with genomics," observes IRRI director general Ron Cantrell.

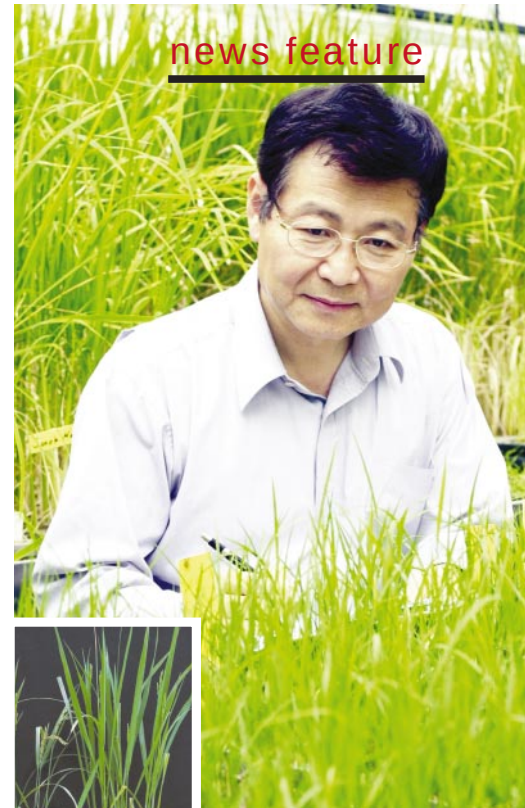
Even if patents do not restrict access to important functional-genomic tools, such resources may not be immediately available

to all researchers. Japan this month launched a programme to assemble a rice genome resource centre at the National Institute of Agrobiological Sciences, but most of its mutants, microarrays and other tools will not be released for a couple of years — and even then, Japanese researchers will get their hands on them first, giving them a head start on work that may prove to be economically valuable. Indeed, most nations with an interest in rice research, including China, are now assembling their own functional-genomics resources to ensure that they are not left at a disadvantage.

Genome duplication

This competition and parallelism of effort is making some experts uneasy. "Not everyone needs a microarray or knockout programme," argues Hei Leung, a plant pathologist at IRRI who is trying to convince key players to combine their resources for the greater good of both research and agriculture. Leung has sown the seeds of an international functional-genomics consortium involving 20 members from 17 institutions around the world, and he aims to form a resource centre that by 2010 will hold mutants for 90% of all rice genes, of which 60% will have been functionally analysed.

Many plant breeders remain sceptical about the value of genomic information, however. The most important contribution so far has been the provision of genetic markers, DNA sequences located near genes that can be tracked to breed for traits that are difficult to observe — a process called 'marker-assisted selection'⁹. There are markers, for example, for more than 20 rice genes that confer resistance to a disease known as bacterial leaf blight. Breeders would like to combine many of these genes for maximum effect, but the effect of a certain gene might be slight in the particular testing environment. The presence of the markers, which



Among the mutants created by Gynheung An (above) is a rice plant (left) that flowers later in the growing season.

can be easily verified, can tell a breeder that a trait gene is present.

But markers can frustrate breeders by becoming detached from the gene when chromosomes recombine during breeding crosses. This is especially true for complex traits, such as drought resistance, in which many genes may be involved. "Markers can be hard to use with confidence," says David Mackill, head of plant breeding at IRRI.

Having more markers would increase the chance of finding one that is tightly linked to the gene for a particular desirable trait. With this in mind, a consortium led by McCouch and researchers at IRRI is aiming to increase the number of markers available.

What about the provision of new information on the function of rice genes? Such knowledge will be welcome, but simply knowing the location of a given gene and the effect of its disruption in a mutant plant is of limited value to plant breeders. They are more interested in knowing what natural variants, or alleles, the gene has — and how to recognize these alleles in crosses. "Too often, gene-discovery studies do not take the work to the resolution that is really needed for applications in breeding," says McCouch.

She is trying to go the extra mile to make genomic information more useful to breeders. At the February meeting in Tsukuba, McCouch described her analysis of the gene that underpinned the 'green revolution' of the 1960s by producing dwarf rice plants that



Full steam ahead: will rice breeders use genomic knowledge to generate new varieties for field tests?



Mister rice guy: genomics could help to ensure that rice distributors' varieties are the genuine article.

put more of their energy into making grain, rather than stems. The precise identity of the gene was finally revealed last year¹⁰, and McCouch has now developed a method that can differentiate between its individual alleles. After listening to McCouch's talk, one excited breeder in the audience said: "That's exactly what we need from the genome."

Tracking traits

Microarrays might also be used to screen for useful new alleles. IRRI alone has collected 100,000 varieties of rice from around the world — most of these are not good crop plants, but they are thought to hold many unknown alleles that might be bred into existing varieties to useful effect. Microarrays spotted with genes that are known to be of agronomic significance could take much of the work — not to mention guesswork — out of identifying potentially valuable alleles. But it would not be cheap — with microarrays costing around US\$400 apiece, a comprehensive screen would be way beyond IRRI's current resources.

In fact, high costs are at the top of breeders' list of objections to functional genomics. "If they gave me a million dollars, I'd put it all into breeding projects," says Glenn Gregorio, a breeder at IRRI who is trying to improve the nutritional value of rice. "Maybe if I had another million, I would put it into functional genomics." But the reality is that

Gurdev Khush: rice breeders still have to find the right gene combinations.

funding for rice research at IRRI and elsewhere is extremely tight.

Sometimes, even knowing an allele and its markers may not be enough. Yano's flowering-time gene, a natural variant, offered breeders an allele with which to control the timing of harvest. But attempts to breed the allele into commercial varieties of rice have proved a big disappointment. "They taste bad," says one breeder at a regional Japanese agricultural station. "It hasn't really been any use."

This example neatly illustrates the difficulty of turning genomic information into agronomic advances. "Gene identification is one thing, but combining the genes as required is another," says Gurdev Khush, now at the University of California, Davis, and formerly of IRRI, where he developed 308 varieties of rice for release during his career. No matter how closely a gene is studied in the lab, it is down to the breeder to discover how it will function in different varieties, or when faced with environmental stresses such as drought and high salinity. "Most molecular biologists don't understand this," says Cantrell.

This gap in understanding is partly down to the fact that, scientifically speaking, genome researchers and breeders are speaking different languages. To address this problem, McCouch has submitted a proposal to the US National Science

Foundation for funding to put together a 'controlled vocabulary' to help breeders, who talk in terms of traits and alleles, communicate more effectively with the geneticists, who focus on genes

and pathways. At present, the two groups even use different yardsticks to measure distances along chromosomes.

IRRI, which in recent years has seen its efforts overshadowed by the publicity generated by rice genomics, hopes to play a key role in bridging the gap between genome researchers and plant breeders. This will be difficult, however, given the diverse demands placed on breeders for varieties of rice that are suited to specific environments — highly productive irrigated areas, lowland areas that receive some rain, flood-prone regions, upland areas, and so on. Depending on the location, breeders might need to incorporate into native varieties traits such as drought resistance, submergence resistance, disease resistance and salinity tolerance.

Mass catering

To meet these diverse needs, IRRI may have to take a step back from trying to deliver finished products, and instead serve up functional-genomics tools, alleles and markers that can be used by local breeders to address their various needs. "We have to supply information," says Cantrell.

Another problem is that, with investment in rice genomics being dominated by Japan and other developed countries, many of the functional-genomic tools now being developed are optimized for their favoured *japonica* subspecies. Inevitably, this will restrict progress in analysing the genome of *indica*, the staple that feeds most of the rice-eating world.

In affluent Japan, the basic agronomic traits that are of primary concern to breeders in developing countries are often of secondary importance to the question of eating quality. Recently, for example, breeders in the Fukui region came up with a set of markers for the gene that gives the desired 'stickiness' to the popular *koshihikari* variety of *japonica*, in the hope that these will help in breeding the trait into other varieties. In a separate endeavour, a company called Plant Genome Center in Tsukuba has developed a set of markers to help to unmask producers who are attempting to pass inferior varieties off as *koshihikari*.

Despite the complexities involved, enthusiasts for rice genomics remain confident that the investment in the rice genome will eventually yield a valuable harvest. But those who were led by some of last year's headlines to expect an imminent agricultural revolution will have to learn to be patient.

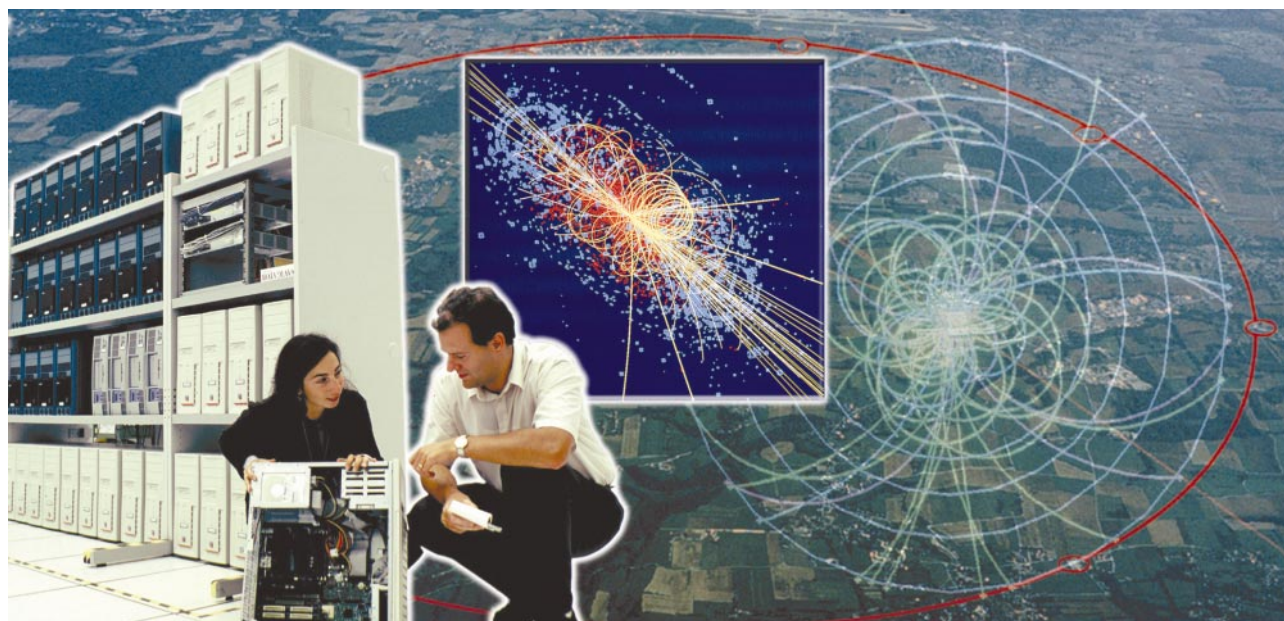
David Cyranoski is Nature's Asian-Pacific correspondent.

1. Yu, J. et al. *Science* **296**, 79–92 (2002).
2. Goff, S. A. et al. *Science* **296**, 92–100 (2002).
3. Sasaki, T. et al. *Nature* **420**, 312–316 (2002).
4. Feng, Q. et al. *Nature* **420**, 316–320 (2002).
5. Yano, M. et al. *Plant Cell* **12**, 2473–2484 (2000).
6. Jeong, D.-H. et al. *Plant Physiol.* **130**, 1636–1644 (2002).
7. Terada, R., Urawa, H., Inagaki, Y., Tsugane, K. & Iida, S. *Nature Biotechnol.* **20**, 1030–1034 (2002).
8. Zhu, T. et al. *Plant Biotechnol. J.* **1**, 59–70 (2003).
9. Knight, J. *Nature* **421**, 568–570 (2003).
10. Sasaki, A. et al. *Nature* **416**, 701–702 (2002).



Tomorrow's computing today

The physics lab that gave us the World Wide Web is now gearing up to make the first practical deployment of the Internet's next big thing: the Grid. Declan Butler takes a test drive.



P. LOIEZ/E. MALANDAIN/CERN

So this is it ... the future of scientific computing. I'm at the headquarters near Geneva of CERN, the European Laboratory for Particle Physics, getting a sneak preview of a prototype Grid. The advanced publicity has led me to expect something special. Grids, according to the hype, will soon allow scientists to harness the power of a global supercomputing network. Just think about it: greater number-crunching power than in your wildest dreams, and seamless access to mountains of raw data stored in colleagues' labs worldwide.

From an ordinary-looking Linux PC, you submit a supercomputing job in high-energy physics. A few hours later, the Grid tells you the results are ready. To be frank, I'm underwhelmed. But that's just what David Williams, formerly head of CERN's computing and networks division, now responsible for the lab's relations with the European Union (EU), wanted to hear. "Grids will have succeeded when they become invisible to the user; when we stop talking about them," he says.

It is only when you peer under the hood that you see why those in the know are getting excited about Grids, and why they may revolutionize the way in which research is conducted. Grids should be able to give scientists

unprecedented computing speed, data access and functionality thanks to a suite of software tools known as 'Globus middleware', which was developed by Grid godfathers Ian Foster of the Argonne National Laboratory near Chicago, and Carl Kesselman of University of Southern California's Information Sciences Institute in Los Angeles.

Live link-up

The smartest component is the 'resource broker'. When you give the Grid a job, this goes off and negotiates with computers signed up to its network to book the computing power you need, which might include a supercomputer in Boston and a 'farm' of thousands of PCs in Madrid. It also scours data centres worldwide for the files and data that you want, and, if needs be, converts them into compatible formats. It will even fire up the software needed to perform a particular analysis.

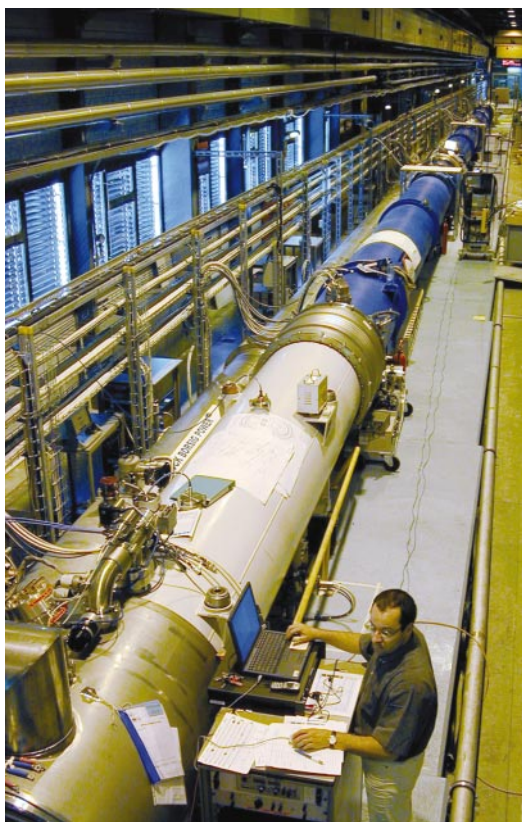
To pull this off, the resource broker must chatter constantly with the other pillars of the Grid middleware. The 'information service' keeps its finger on the pulse of the network, tracking up-to-the-minute information on where and how much computing power is available, and how fast various data links are running. The 'replica catalogue' is

Combined effort: physicists hoping to identify the signatures of rare subatomic particles (inset) plan to link up computers around the world to give them the data-crunching power they need.

the Grid's Yellow Pages, keeping a note of the type and whereabouts of all data and files. The 'logging and bookkeeping' server, meanwhile, acts as an office manager, tracking the status and history of all submitted jobs.

Many labs are working on prototype Grids, but at CERN the concept is about to be given its first proper outing. Here, the geeks are working with a gun against their heads. They have been set a deadline of July to prove that the Grid can deliver to the lab's scientists the awesome computing power that will be needed when the US\$2-billion Large Hadron Collider (LHC) comes online in 2007.

By smashing protons into one another at close to the speed of light, the LHC will spew out petabytes (10^{15} bytes) of data. That's about the same quantity of information involved if everyone on Earth were each talking into 20 telephones simultaneously. Clever algorithms that filter out 'noise', events that physicists are reasonably confident are not interesting, will eventually stem the flow somewhat. But if researchers across more than 40 countries are to detect the signatures of exotic subatomic



Information overload: the Large Hadron Collider will generate huge amounts of data.

particles such as the Higgs boson, which is hypothesized to give other particles their mass, only the Grid will do. "It's not an exaggeration to say: 'no Grid, no LHC,'" says Williams.

Given these high stakes, high-energy physicists want to do a few 'dry runs' — which is where the July deadline comes in. CERN is now gearing up for a groundbreaking project in which physicists will simulate the creation and analysis of the giant sets of data that they will have to handle, and so 'tune' their computing models and software. "This is not yet another Grid technology project, it is a Grid deployment project," says Les Robertson, who is heading the effort.

Until now, CERN has been operating the EU-funded European DataGrid testbed. This involves some 1,000 networked machines in 15 countries, 5 terabytes (5×10^{12} bytes) of disk space and 350 test users. July is about making this network available to all LHC physicists. From there on comes a series of deadlines by which the network must be ratcheted up to a system that will eventually require the equivalent of 200,000 of today's fastest PCs, store up to 8 petabytes of data annually — the same as 16 million CD-ROMs — and serve 8,000 physicists across the world.

The challenge facing CERN's Grid specialists is the problem of 'scalability', the catch-all term used by engineers to explain why something that worked in the prototype doesn't work in the real world. Take the resource

broker. When physicists run operations such as Monte Carlo simulations — a common statistical technique that can create many computing tasks at once — it can get hammered with requests and often packs up. To get round this, information technology (IT) staff at CERN are rewriting their software to include multiple resource brokers in the Grid working in concert to distribute the load.

Although Grids are designed to let you weave together myriad computers or operating systems into one seamless whole, and shuffle data around irrespective of location, in practice incompatibilities crop up all of the time. "There is a gap between the hype and technical performance," admits Ben Segal, who formally retired last year from his post as manager of the European DataGrid, but still spends some of his time working on the project.

Language barriers

Although scientists worldwide know how to collaborate, their computers do not. Compilers, for instance, are key components of any computer that convert the various languages used to write software into machine code that a computer can understand and execute. But the compilers used in LHC experiments often can't communicate with those on machines around CERN's Grid, which likewise are often unable to talk to one another. Quality of service is another major problem: one university in the Grid may have an excellent IT infrastructure, but another's systems may crash all the time.

Add these problems together, and it becomes clear why 20% of all jobs submitted to the prototype LHC Grid currently fail to complete. At present, keeping the network online also requires a great degree of human intervention. "You've got to turn all the knobs simultaneously to the right place and hold them there if it is to work," says Ian Bird,



Ian Bird is confident the compatibility problems that may limit the Grid can be overcome.

deployment manager for the LHC Grid.

But if anyone can succeed in ironing out the glitches, CERN's computing specialists would be among the favourites. To avoid the Grid descending into gridlock, a key step over the coming months will be to marry the best technologies from CERN's own European DataGrid with three parallel research efforts in the United States: the Particle Physics Data Grid, the Grid Physics Network and the International Virtual Data Grid Laboratory. It won't be easy, but Bird is quietly confident. "Of course we will make it," he chuckles.

In addition to the technical obstacles, getting the LHC Grid online requires unprecedented formal collaboration between the IT service departments of the institutions taking part, to agree on joint standards and best practices. Any institution taking part in the Grid has, for example, the right to carry out on-site inspections to verify that another's security and system performance are up to scratch. Thankfully, says Roger Jones, a physicist at Lancaster University, UK, it has proved "surprisingly easy" to get agreement among institutions, probably because of the high stakes involved.

Indeed, one could argue that the human element of Grids is just as important as the fancy technology. Increasingly, argue Grid enthusiasts, scientists will see themselves less as belonging to individual bricks-and-mortar institutions, and more as members of 'virtual organizations', communities of researchers in defined research areas or associated with particular experiments, who together decide what computing and data resources they will share over the Grid. Gone will be the logjams caused by limitations in computing power and data storage at one institution, and the need to rack up endless frequent-flyer miles to participate effectively in a project.

The primary goal of the LHC Computing Grid is to prove that physicists will be able to handle the avalanche of data that they will be presented with when CERN's new collider comes online, and to make all this information seamlessly available to individual researchers as if it were all on their own hard disk. But the project is also an experiment in a new sociology of e-science. CERN hopes to show the way, and encourage other disciplines to embrace the Grid's possibilities.

That may sound like a grandiose goal, but remember that CERN has already once changed the way in which the world interacts with computers, through the work of Tim Berners-Lee, who invented the technology that became the World Wide Web as a means of allowing physicists to share documents. Turning the Grid from testbed to reality would be a fitting encore. ■

Declan Butler is Nature's European correspondent.

LHC Computing Grid

♦ <http://lhc.web.cern.ch/LCG>

European DataGrid

♦ <http://eu-datagrid.web.cern.ch/eu-datagrid>

Unrestricted free access works and must continue

Bioinformatics researchers shouldn't need coercion to act responsibly and collegially.

Sir — The policy on release of unpublished data from large genome centres has generated considerable discussion and some confusion, as your Editorial “Sacrifice for the greater good?” makes plain (*Nature* **421**, 875; 2003). In our view, data sets from large, centralized, expensive genome data-collection projects should be freely available to the entire scientific community, immediately and with no restrictions or conditions.

Our position is that pre-publication release of large genome data sets is a special case, and not a principle that should be applied “throughout the world of biology”, as was asserted in your editorial. Large genome sequencing has become an expensive, factory-style operation, in which economies of scale can only be realized by very large centres. Large data production centres, established and supported by the scientific community, represent a different model of science from traditional investigator-initiated projects. We argue that they need to operate under different rules.

The broader scientific community supports the highly centralized model represented by the US large-scale genome centres (funded via the National Institutes of Health and the Department of Energy) on the condition that everyone in this community gets equal access to the data. If this is the case, everyone wins: large data sets are generated at the lowest cost and greatest speed, and scientific work progresses on multiple fronts without delay. In contrast, if genome centres restrict their data and get preferential access to it, then some members of the community will no longer support monopolistic funding models (in which large centres sequence one genome after another without peer review of each project). Instead, they will demand the right to compete with these empires, especially for the most scientifically desirable genomes. Other scientists, especially bioinformaticians, will seek to relocate to the centres to gain the advantage of early data access. Data restrictions will therefore promote factionalization where we should be seeking efficiencies of scale, and centralization where we should be promoting diversity.

We agree with your editorial that the proposed new policy, recently released for comment by the US genome centres (see www.genome.gov/page.cfm?pageID=10506537), is ambiguous. It states that genome sequence data “should be available

for all to use without restriction”. This statement, which notably uses the word “should” rather than “must”, is qualified by a lengthy discussion of conditions, including a reference to the sequence producers’ interest in “the first peer-reviewed published analysis of the results of the sequencing project”. This reflects the real concern of the genome centres that prepublication data access may put the scientists there at a competitive disadvantage. We understand these concerns, but we believe that the qualifying discussion cannot coexist with the principle of “without restriction”. We propose that these qualifications are simply dropped to avoid confusion among data consumers, journals and journal reviewers. The Human Genome Project has been a spectacularly successful demonstration that the “Bermuda rules” of free access without restriction do work, for everyone.

As bioinformaticians, we have an important role in this process. We reaffirm our own commitment to respecting the

goals of the scientists at the genome centres, who should be consulted as part of any large-scale analysis of unpublished genome data, and included as collaborators where appropriate. It is a serious problem that these invaluable centres feel compelled to coerce such simple scientific courtesy from our community. We encourage our peers in bioinformatics to act responsibly, cooperatively and collegially, to help to assure open, unrestricted, immediate release from large community-driven data-collection projects.

Steven Salzberg*, Ewan Birney†, Sean Eddy‡, Owen White*

*The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA
salzberg@tigr.org

†European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK

‡Howard Hughes Medical Institute and Department of Genetics, Washington University School of Medicine, 4566 Scott Avenue, St Louis, Missouri 63110, USA

Watching an IT icon slide into history

Sir — In the famous photographs of James Watson and Francis Crick with their DNA model (reproduced in *Nature* **421**, 15 & 417; 2003) a small icon of information technology can be seen in Crick's right hand — a slide rule. (One wonders, incidentally, why this was shown in the photograph instead of the “measuring stick” which, together with a plumb line, they used “to obtain the relative positions of all atoms”, as described by Watson in *The Double Helix* — anecdote has it that Crick told the



Ruled out: unlike Francis Crick, it seems that the slide rule is slowly being forgotten.

insistent photographer that the slide rule was irrelevant.)

At Newcastle University we have been teaching biology students the use of computers for nearly three decades, beginning each year with a brief account of the staggering advances in computational power since the middle of the twentieth century. Each year, we show them an example of what Crick held. It seems fitting, on the fiftieth anniversary of the reconstruction of DNA, and at the advent of the rise of computational biology (C. Surridge *Nature* **420**, 205; 2002), to report that this year, for the first time, none of our first-year students recognizes a slide rule.

W. John Cram

School of Biology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK

Secrecy is increasing in step with competition

Sir — Various authors and reports have recently claimed (for example, refs 1–3) that the increasing commercialization of academic science has led to an increase in secrecy. However, our comparison of two surveys of experimental biologists, mathematicians and physicists, conducted about 30 years apart, suggests a more

complicated and interesting picture.

As feared, secrecy (measured as unwillingness to discuss ongoing research with those outside the research group) has increased.

In 1966 (ref. 4), 50% of 1,042 respondents reported feeling safe in talking with all others about their current research, but by 1998, when we surveyed 202 scientists from the same three fields (details of methods and results available from J.W.), the equivalent number was 26%. Experimental biologists have become particularly secretive, with only 14% willing to talk openly about their current research in 1998. Secrecy is strongly predicted by scientific competition (measured as concern over having one's research results anticipated). The effects of commercial activity, on the other hand, are quite mixed. Patenting has no effect; industry funding is associated with greater secrecy; but having industry collaborators is associated with less secrecy.

These university–industry collaborations can be viewed as part of a strategy to share findings and expertise with the wider scientific and technical community. For companies, timeliness and customization of information are often more important than exclusivity, so they are willing to tolerate, even encourage, their academic collaborators' participation in the shared conversation of a scientific field, thereby giving the company access to the whole community's expertise. In contrast to these collaborations, industry funding alone is often associated with a university laboratory acting as a subcontractor to a company's R&D project, and may produce associated secretive behaviour.

Thus, there is reason to believe that secrecy has increased among academic scientists, but that the focus on commercialization as the cause may be misplaced. Although commercial activity may reduce formal activities such as publication or sharing of materials, it may have fewer negative effects on informal communication among researchers. As this informal communication is significant in transferring information to companies⁵ and is at least as important as publication for distributing information among scientists, this is encouraging news.

Although it is right to raise concerns about the negative effects of publication restrictions associated with industry funding, we should not conclude that university–industry linkages per se produce unhealthy levels of secretiveness among academic scientists. Instead, it may be better to focus on alleviating some of the negative consequences of scientific competition.

Recent increases in US government funding for science, if they are sustained, may help to lower the intensity of competition, as well as the dependence on industry funding, and thereby reduce secrecy. Furthermore, although we need to be wary of the strings attached to industry funding, university–industry collaborative research should be encouraged.

We thank Lowell Hargens for providing field-level data from ref. 4.

John P. Walsh*†, Wei Hong†

*Research Center for Advanced Economic Engineering, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
walsh@aee.u-tokyo.ac.jp.

†Department of Sociology, University of Illinois, Chicago, Illinois 60607-7140 USA

1. Blumenthal, D. *et al.* *J. Am. Med. Assoc.* **277**, 1224–1228 (1997).
2. Cook-Deegan, R. M. & McCormack, S. J. *Science* **293**, 217 (2001).
3. Campbell, E. G. *et al.* *J. Am. Med. Assoc.* **287**, 473–480 (2002).
4. Hagstrom, W. O. *Am. Sociol. Rev.* **39**, 1–18 (1974).
5. Cohen, W. M., Nelson, R. R. & Walsh, J. P. *Management Sci.* **48**, 1–23 (2002).

Pre-genomics training hinders Indian biotech

Sir — Ashok Parthasarathi in his Commentary “India: a champion of new technologies” (ref. 1) rightly affirms India's bid to play a leading part in global technology developments. India's technological independence has come mainly in the physical sciences, such as space research, telecommunications, software, defence systems, energy and supercomputing.

Although Parthasarathi lists biotechnology as a key ‘champion technology’, a tremendous boost is needed for it to make a global impact.

India shares this predicament with other developing countries that have a vast research force but the training and infrastructure of pre-genomics days, and little experience of the research skills required for post-genomic biology. With the genome sequence data for several organisms, including humans, now available, the rules of knowledge-based commercial ventures have changed and an era of ‘omics’ — genomics, proteomics and so on — has emerged², with functional genomics as the new ‘mantra’.

In order to become a champion of biotechnology, India needs a paradigm shift in the organization of research and training, priming research institutions and universities to change gear and meet global challenges. Two aspects need urgent attention: establishing functional-genomics centres for biotechnology and

related basic science research, as is done in China³; and training researchers in state-of-the-art skills, along the lines of initiatives by the European Molecular Biology Laboratory.

All the stakeholders — in the fields of policy, administration, science and industry — have to address this problem and give a directional nudge to research initiatives.

Nature has covered some of these issues recently, for example the low impact factors of Indian biological-science journals⁴; the urgent need to explore the best scientific options for sustainable development by regional centres of the International Council of Science⁵; and the effect of post-genomics research on traditional methods of food production⁶. Informal articles on topics such as these serve as a compass for framing policies and making course corrections during implementation.

R. M. Ranganath

Department of Botany, Bangalore University, Jnanabharathi Campus, Bangalore 560056, India

1. Parthasarathi, A. *Nature* **422**, 17–18 (2003).
2. Palsson, B. *Nature Biotechnol.* **20**, 649–650 (2002).
3. Cyranoski, D. *Nature* **410**, 10–12 (2001).
4. Vohora, S. B. & Vohora, D. *Nature* **412**, 583 (2001).
5. Clarke, T. *Nature* **420**, 733 (2002).
6. Knight, J. *Nature* **421**, 568–570 (2003).

Should tobacco ban rule out governments too?

Sir — I was interested to read in your News story “Academics fume as university refuses to reject tobacco dollars” (*Nature* **422**, 361; 2003) a scientist at the University of California, San Francisco, quoted as saying it is immoral to accept research funds from tobacco companies because “it is not appropriate to take money from an industry that kills 5 million people worldwide and constantly lies”.

I guess that means that no research funds should be accepted from the US Department of Defense, nor from its corporate headquarters — the US government. Other national governments would presumably also be banned from supporting research under these new politically correct guidelines.

Richard S. Jope

Department of Psychiatry, Sparks Center 1057, 1720 7th Avenue South, University of Alabama at Birmingham, Birmingham, Alabama 35294-0017, USA

correspondence

Contributions to Correspondence may be submitted to corres@nature.com. They should be no longer than 500 words, and ideally shorter. Published contributions are edited.

Who cares about the double helix?

Collective memory links the past to the future in science as well as history.

Bruno J. Strasser

Science is about building the future. Discovering new mechanisms that operate in nature, inventing new tools of investigation and elaborating new concepts are the core of scientific practice. Then why do past scientific events, such as the discovery of the DNA double helix¹, still receive attention 50 years after they took place? Are they simply scientific folklore, and their commemoration merely entertaining rituals? Or is it that our collective memories of the past play a crucial contemporary role by sustaining experimental designs, research programmes, funding channels and collective identities? In other words, when we refer back to a discovery today, it is because we want to say something about the present?

It is too early to assess the deeper meaning of this year's DNA double-helix commemorations — how they will shape our collective memories of the discovery and for what current agendas they are being used. A critical look at these rituals as they take place will allow us to witness collective memory in the making. With the history of the double helix in mind², we can learn something about the values and ideals that make science what it is today.

We usually think that the double-helix model acquired immediate and enduring success. On the contrary, it enjoyed only a “quiet debut”³. Only when the role of DNA in protein synthesis became clearer, in the late 1950s, did biochemists, for example, take a serious interest in it. Even in Cambridge, UK, where the structure was discovered, it seems to have had “only a subordinate role” in negotiations over the future of the laboratory in the late 1950s⁴. So when did it become so immensely popular in the scientific community? The 1960s? 1970s? 1980s? None of these. Not until the 1990s (Fig. 1).

Prizes and popularity

True, when James Watson and Francis Crick (with Maurice Wilkins) were awarded the Nobel prize in 1962, their seminal work received increased attention the following year (see Fig. 1) which coincided with the tenth anniversary of their discovery. However, after that time, the 1953 paper followed the familiar declining citation pattern. Even Watson's bestselling book, *The Double Helix* (1968), did not stimulate citations of the original paper, perhaps because of the scandal it provoked in some scientific circles. Yet in the early 1990s, something changed — citations began to rise dramatically to the point where 2003 promises to be an all-time record. Of course, the double helix achieved

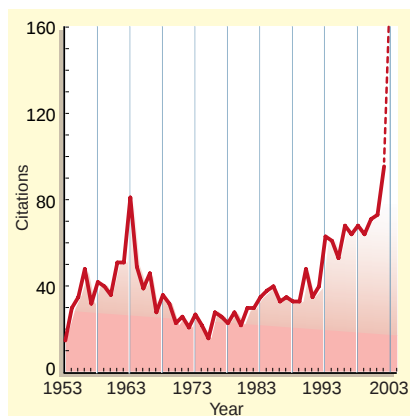


Figure 1 Citations per year of Watson and Crick's first 1953 *Nature* paper in scientific journals (English language only). The value for 2003 is extrapolated linearly from data available as of 31 March 2003 (40 citations). Source: ISI Science Citation Index.

iconic status well before the 1990s, attracting many artists during the past 50 years. However, of the numerous examples used in a recent account of this “Mona Lisa of modern science”⁵, only one is before the 1990s. DNA seems to have become a central theme in contemporary art since only the 1990s⁶.

In scientific circles, the double helix did much in the late 1950s and 1960s to bring together crystallographers, biochemists and phage geneticists around a new identity — ‘molecular biology’, as they called it. This new social grouping could take place around the helix as it was a discovery to which the different communities believed they had contributed. The DNA double helix became a ‘totem’ for the widely diverse tribe of molecular biologists. The retrospective account of the discovery as being at the origin of molecular biology was constructed in the 1960s by molecular biologists themselves to promote their discipline⁷.

Why did the popularity not only of the double helix, but of Watson and Crick's discovery of it, escalate at the end of the twentieth century? Citing Watson and Crick's first 1953 paper contributes to the construction of collective memory — social groups' shared representations of their past. Historians now pay much attention to how collective memories (of the Holocaust and of the French Revolution, for example) shape identity, and how they are constructed and transmitted through commemoration and oral tradition. It is not surprising that a similar process exists in science⁸.

Collective memory is directed towards

the past as much as towards the present. Narratives about past science can be used to illustrate or justify existing modes of research organization and specific experimental approaches — sometimes even bringing them into being. For example, the collective memory of penicillin, the wonder drug of the Second World War, has promoted a specific way of thinking about therapy and a specific manner of conducting therapeutic research.

Similarly, Linus Pauling's famous 1949 paper on sickle-cell anaemia became a citation classic⁹ very much in the same way as Watson and Crick's. Over the past 50 years, Pauling's research on sickle-cell anaemia has been incorporated in very different — often incompatible — narratives constructed by diverse scientific communities. In its most recent version, the collective memory of sickle-cell anaemia as a ‘molecular disease’ has sustained a particular kind of therapeutic research, targeting the haemoglobin molecule itself or the gene responsible, instead of some other step along the chain of events that lead to the pathological consequences affecting the patient. Citing Pauling's work is an efficient way to obtain justification to counter sceptics of gene therapy. In this sense, not only does history become transformed into memory, but memory makes history. Collective memory links the past with the future.

Linking back

To understand the construction of our collective memory of the double-helix discovery we need to know who is citing Watson and Crick's paper today, in what context and why. If we exclude the strictly commemorative citations of this year — more on these later — most citations are to be found in editorials and the discussions of research papers. As it does not convey any new information to the reader to mention that DNA has a double-helix structure, reference to it must serve different purposes.

One of these is to construct genealogies, a long-standing favourite way to draw boundaries around social (academic) territories⁷. Authors in an unprecedented variety of fields, from neurosurgery to plant physiology, cite the 1953 paper in an attempt to appropriate the current prestige of molecular biology for their own speciality by underlining (sometimes far-fetched) historical affiliations with the double helix. “The early 1950s,” wrote Edward Wood in 2001, “saw biochemistry give birth to molecular biology. James Watson and Francis Crick's paper on the molecular structure of nucleic acids was published.”¹⁰ The fact that Watson and Crick knew very



Tribal totem: the DNA double helix serves as a unifying symbol and provides a collective identity for many different areas of science.

little biochemistry when they built their DNA model gives a curious twist to this argument. For David Miska, working in a biotech-pharmaceutical consultancy: "The intellectual parent of the biotechnology industry has its own birthday in April 2003, which marks the 50th anniversary of the epochal publication of the structure of DNA by Watson and Crick."¹¹ In a broad sense, the biotechnology industry largely predated 1953. In a narrower sense, as a DNA-based industry, it owes more to the discovery of restriction enzymes and recombinant DNA technology than to the discovery of the molecule's double-helical structure. The authors of these accounts are not professional historians, and such historical inaccuracies are relatively benign. There are more interesting avenues to be explored.

The key to understanding the renewed fame of the double helix at the end of the twentieth century is the Human Genome Project (HGP), which began in 1990. It brought the double helix, and Watson and Crick's discovery of it, to centre stage as never before. The HGP, of course, has an inherent if intangible link with the earlier discovery, namely the fact that DNA is the central object being investigated. Other events have reinforced this association — James Watson was the first director of the HGP; the HGP was a US–UK effort (Watson and Crick were a US–UK team); and the HGP was supported by US and UK funding agencies.

The promoters of the HGP have gone further in strengthening their association with the double helix. For example, in 1998 Francis Collins, the HGP's subsequent director, announced the project's goal: "Finish the complete human genome sequence by the end of 2003. The year 2003 is the 50th anniversary

of the discovery of the double helix structure of DNA by James Watson and Francis Crick. There could hardly be a more fitting tribute to this momentous event in biology"¹². When the draft sequence was published in *Nature* in February 2001, *Science's* short history of the HGP, which accompanied its publication of the draft sequence produced by Celera Genomics of Rockville, Maryland, began by mentioning Watson and Crick's 1953 discovery¹³. It soon became common to read assertions such as: "The most significant event in biology since the 1953 publication of the Watson and Crick paper describing the structure of DNA was the publication in *Science* and *Nature* earlier this year of the sequence of the human genome."¹⁴

This deliberate comparison of the HGP with the double-helix discovery by participants and supporters of the project helped to lift its scientific stature above criticisms that it was an unimaginative piece of applied science. Setting the double helix as the highest standard of scientific creativity¹⁵ helped to make the descriptive genome sequencing project look less tedious. After all, both were attempts at describing the structure of DNA and would result only in a guess at the possible "genetic implications", as Watson and Crick had put it. For the

Setting the double helix as the highest standard of scientific creativity helped the HGP look less tedious.

promoters of the genome project, the discovery of the double helix provided a very useful cultural bridge towards public acceptance and identification of its goals. The association reinforced the image of the public consortium project as a piece of fundamental research, above industrial interests and patent disputes, unlike the similar project carried out by its competitor, Celera Genomics. The double helix worked as a marker of an age of (lost) innocence, when youth, intelligence and self-assurance were sufficient to make great discoveries in science. For many, Watson and Crick not only provided a model of DNA, but also a model of scientific practice, where creativity and hard evidence share in the construction of scientific knowledge. ■

Bruno J. Strasser is at the Institute for the History of Medicine and Health (IURHMS), University of Geneva, CMU, CH-1211 Geneva 4, Switzerland.

1. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737–738 (1953).
2. Morange, M. *A History of Molecular Biology* (Harvard Univ. Press, Cambridge, 1998).
3. Olby, R. *Nature* **421**, 402–405 (2003).
4. de Chadarevian, S. *Designs for Life. Molecular Biology after World War II* (Cambridge Univ. Press, Cambridge, 2002).
5. Kemp, M. *Nature* **421**, 416–420 (2003).
6. Anker, S. & Nelkin, D. *The Molecular Gaze: Art in the Genetic Age* (Cold Spring Harbor Laboratory Press, New York, in press 2003).
7. Abir-Am, P. *History of Science* **23**, 73–117 (1985).
8. Abir-Am, P. G. & Elliott, C. A. (eds.) *Commemorative Practices in Science. Historical Perspectives on the Politics of Collective Memory* (Univ. Chicago Press, Chicago, 2000).
9. Strasser, B. *J. Am. J. Med. Genet.* **115**, 83–93 (2002).
10. Wood, E. J. *Nature Rev. Mol. Cell Biol.* **2**, 217–221 (2001).
11. Miska, D. *Nature Rev. Drug Disc.* **2**, 231–233 (2003).
12. Collins, F. S. et al. *Science* **282**, 682–689 (1998).
13. Roberts, L. et al. *Science* **291**, 1195 (2001).
14. Harris, T. *Trends Mol. Med.* **7**, 492–493 (2001).
15. *Nature Genet.* **27**, 227–228 (2001).

Acknowledgements

I thank Soraya de Chadarevian, Jean-Paul Gaudillière, Marc Geiser and Michel Morange for discussions.

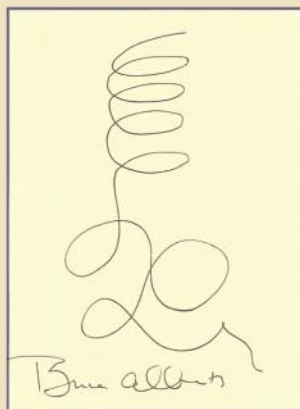
DNA from Alberts to Zinder

Peter Sherwood

To mark the fiftieth anniversary of the discovery of the double helix, a conference was held at Cold Spring Harbor Laboratory ('The Biology of DNA', 26 February to 2 March 2003). Several conference participants were asked to draw their conception of DNA, and to sign and date their sketch 28 February 2003 (50 years since the discovery of base pairing). A sample of these drawings appears here.

The drawings span many topics in the biology of DNA, including genetics, the cell cycle, replication, recombination, repair, transposition, topology, telomeres, forensics, genomics, evolution and Z-DNA. Drawings that use DNA as a springboard to illustrate RNA splicing and ribozymes are also featured. The full collection can be seen at www.nature.com/nature/DNA50.

Peter Sherwood is at Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, New York 11724, USA.
e-mail: sherwood@cshl.edu



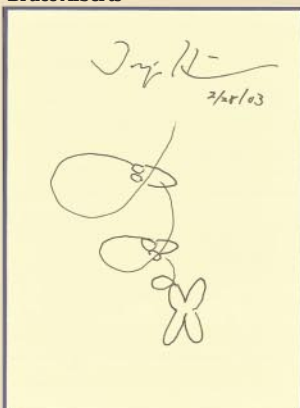
Bruce Alberts



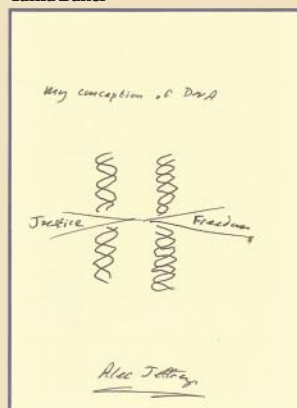
Tania Baker



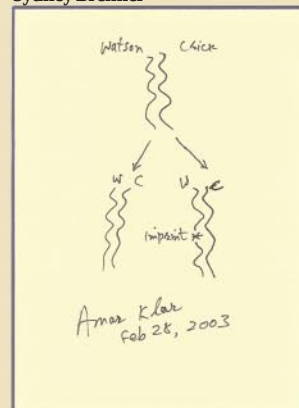
Sydney Brenner



Tatsuya Hirano



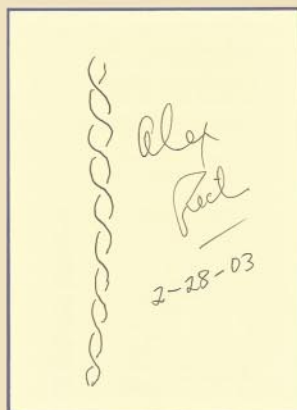
Alec Jeffreys



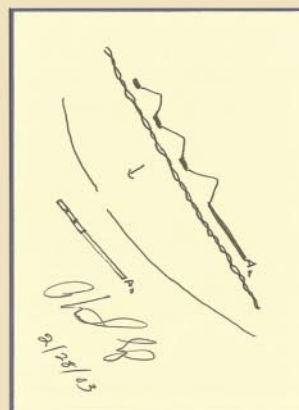
Amar Klar



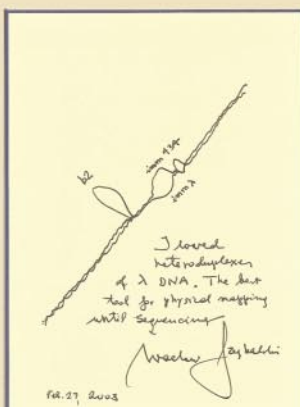
Arthur Pardee



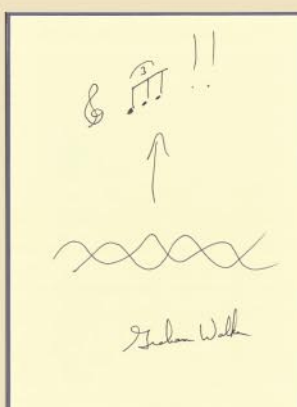
Alex Rich



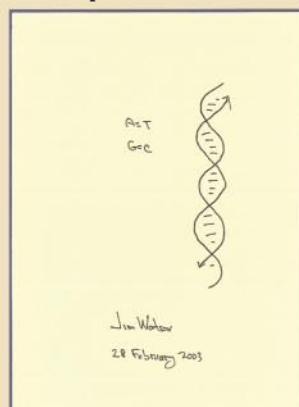
Phil Sharp



Waclaw Szybalski



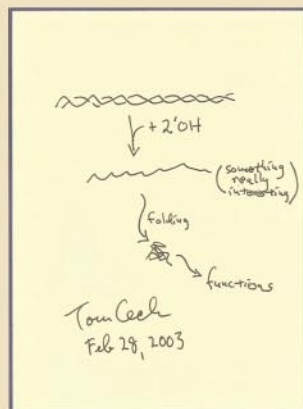
Graham Walker



James Watson



Mario Capecchi



Tom Cech



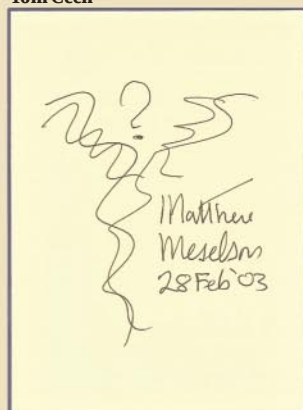
Francis Collins



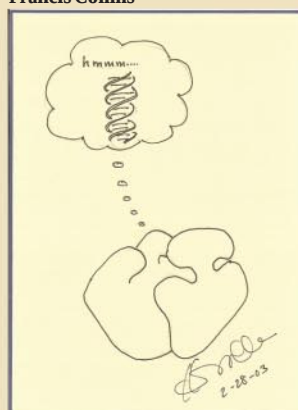
Walter Gilbert



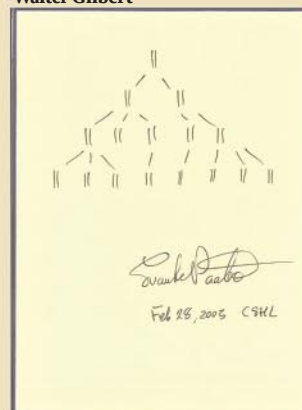
John Kuriyan



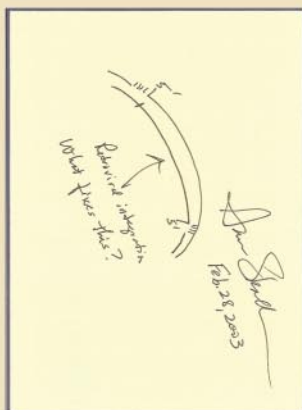
Matthew Meselson



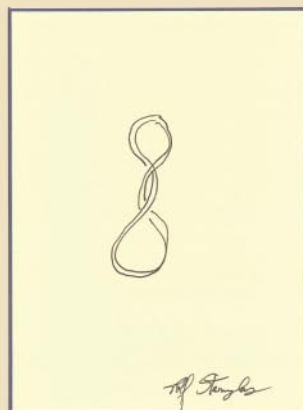
Harry Noller



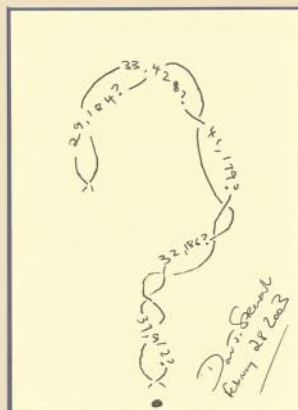
Svante Pääbo



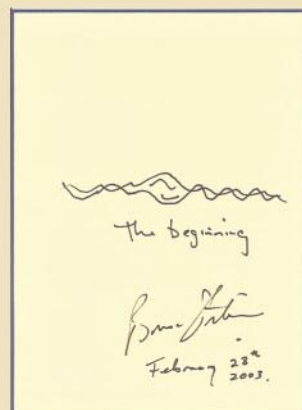
Anna Marie Skalka



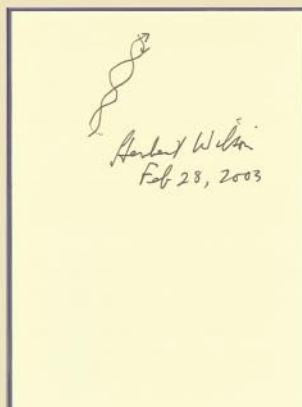
Rolf Sternglanz



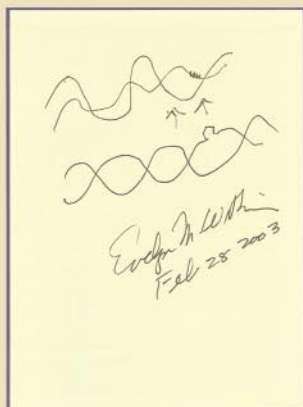
David Stewart



Bruce Stillman



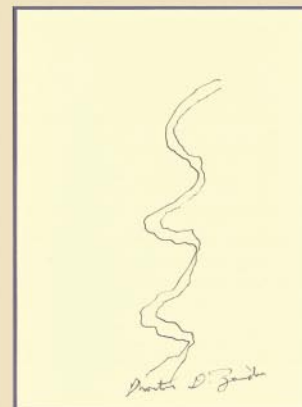
Herbert Wilson



Evelyn Witkin



Jan Witkowski



Norton Zinder

Life, the movie

Fifty years after revealing the structure of DNA, James Watson looks back.

DNA: The Secret of Life

by James D. Watson with Andrew Berry

Alfred A. Knopf: 2003. 416 pp. \$29.95.

William Heinemann: 2003. £20

Maxine Singer

In an age when celebrities matter, biology has produced two: the DNA double helix and James D. Watson. Both emerged 50 years ago, and celebrations of that anniversary, including this book, abound. Lest he be forgotten, the book is dedicated to Francis Crick, who, since discovering with Watson the structure of DNA, has eschewed iconic status while contributing important, insightful ideas to biology. In the fashion of movie stars and sports heroes, this book was written 'with' Andrew Berry, who was initially recruited to work on the associated five-part television series, already shown on Channel 4 in Britain and to be aired on the Public Broadcasting System in the United States in April. According to the authors' note, three additional people wrote substantial parts of the book. Perhaps that's why, despite the first-person-singular pretence, I missed Watson's provocative and sometimes maddeningly idiosyncratic style.

All of the people and stories associated with the landmark discoveries in genetics are here, from Gregor Mendel to the latest failures of gene therapy. They include the story of Thomas Hunt Morgan and the chromosomal theory of heredity; George Beadle and the 'one gene—one enzyme' concept; Oswald Avery, Alfred Hershey and Martha Chase and the identification of DNA as the hereditary molecule; François Jacob and Jacques Monod and the regulation of gene expression; Marshall Nirenberg and Gobind Khorana and the genetic code; and Mary-Claire King and the identification of genes associated with breast cancer. Research is realistically portrayed as a human endeavour that is subject to the quirks of personalities.

Unfortunately, however, for all its marvellous and comprehensive sweep, this book is flawed by some distressing errors. For example, in *The Path to the Double Helix*, Robert Olby carefully documents that Crick proposed the antiparallel configuration of the two DNA chains before Watson got the base pairing correct, but here the two insights are reversed. The technique used to produce Affymetrix DNA chips is described as "precise micropipetting" rather than an inspired marriage of computerized photolithography and photochemistry. The history of the development of recombinant-DNA concepts is garbled: combinations of eukaryote and bacterial plasmid or bacteriophage DNA

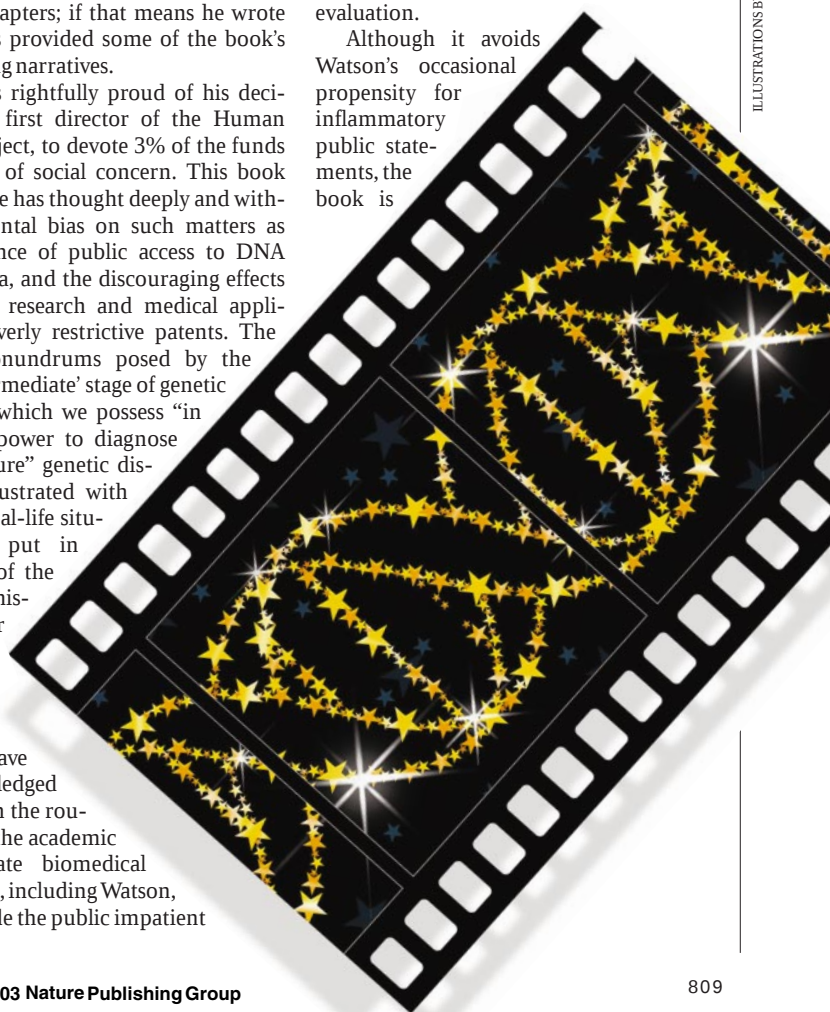
were made by several investigators at Stanford University in California, albeit by inconvenient methods, well before Stanley Cohen and Herbert Boyer devised a convenient way to produce recombinant DNA molecules. The modern form of the method for importing plasmids into *Escherichia coli* bacteria was described by Mandel and Higa in 1970, rather than by Cohen in 1971. In addition, the definition of 'autosomal' is inexplicably confused with that of 'dominant', and the definition of 'cloning' is incomplete.

Genetics has been a lightning rod for public concern for almost a century. Watson and his co-authors recount the issues from the destructive influence of eugenics, through Trofim Lysenko and the controversies over genetically modified plants, to the forensic analysis of DNA. A chapter called "Defying Disease" superbly parses the complex questions surrounding genetic testing and diagnosis in a manner that is informative for biomedical scientists and patients alike. The authors' note acknowledges Jan Witkowski for "pulling together" this and two other chapters; if that means he wrote them, he has provided some of the book's most engaging narratives.

Watson is rightfully proud of his decision, as the first director of the Human Genome Project, to devote 3% of the funds to questions of social concern. This book reveals that he has thought deeply and without judgemental bias on such matters as the importance of public access to DNA sequence data, and the discouraging effects on scientific research and medical applications of overly restrictive patents. The troubling conundrums posed by the present 'intermediate' stage of genetic research, in which we possess "in general the power to diagnose but not to cure" genetic diseases, are illustrated with wrenching real-life situations and put in the context of the analogous history of earlier medical advances. This point is welcome but it could have been acknowledged that too often the routine hype of the academic and corporate biomedical communities, including Watson, has itself made the public impatient for cures.

Curiously, while calling for strict regulation to avoid the dangers of gene therapy, Watson still defends his waffling over the scientific community's actions regarding recombinant-DNA experiments. Although a signatory of the 1974 moratorium letter, he misleadingly describes its contents as calling on scientists worldwide to suspend "all recombinant studies". The letter said no such thing: it asked for a voluntary deferring of only several specific types of recombinant experiment that could "prove biologically hazardous". At the Asilomar Conference in early 1975, six months after the letter was published, Watson joined those who believed that "the moratorium be damned, let's get on with the science". Since then, and again in this book, he has claimed that the scientists' responsible effort yielded "nothing more than five sad years of delay in important research". But the avoidance of restrictive federal legislation in the United States, and this book's description of the rapid scientific progress during the late 1970s, belies his intimidating evaluation.

Although it avoids Watson's occasional propensity for inflammatory public statements, the book is



ILLUSTRATIONS BY DAVID NEWTON

straightforward in decrying the political influence in the United States of fundamentalist religious tenets on the teaching of evolution and the pursuit and application of research on embryos. Watson passionately objects to the current situation, in which US “politicians continue to pander to the outspoken religious minority”. And he speaks for many scientists when he says: “I do not dispute the right of individuals to look to religion for a private moral compass, but I do object to the assumption of too many religious people that atheists live in a moral vacuum.” These statements would be even more powerful if the chapter on recombinant DNA was not entitled “Playing God”, thereby perpetuating a tool used by those who would inhibit modern genetics research.

The unfortunate errors aside, this book is more inclusive and is better reading than similar attempts by science journalists. Although aimed at a general audience, the ignorance among many biomedical scientists of the history of their science suggests that they too can learn a great deal from the book, and enjoy doing so. ■

Maxine Singer is president emerita of the Carnegie Institution of Washington, 1530 P Street Northwest, Washington, DC, 20005, USA.

A final call for peace

The Hedgehog, the Fox, and the Magister's Pox: Mending the Gap Between Science and the Humanities

by Stephen Jay Gould

Harmony: 2003. 288 pp. \$25.95

Published in the UK on 29 May by Jonathan Cape, £18.99

David Hull

Stephen Jay Gould's *Rocks of Ages* investigated the relationship between science and religion. In *The Hedgehog, the Fox, and the Magister's Pox*, published posthumously, Gould turns his attention to the relationship between science and the humanities. During the Renaissance, what we now term ‘scientists’ were at odds with humanists, whose goal was to recover the wisdom of antiquity, not to generate new ideas by means of empirical investigation. An appropriate mantra for Renaissance humanist scholars was ‘been there, done that’. When the sciences cast off the heavy hand of the humanities in the Renaissance, a new adversarial attitude took its place. In this book, Gould strives to outline a more peaceful, mutually supportive view of the relationship between the sciences and the humanities.

Gould sees three ways in which the humanities can contribute to science. First, “science needs the humanities to teach us the quirky and richly subjective side of our

own enterprise”. The path of science does not run smoothly — at times it is indeed quirky. But is science really subjective, even richly subjective? Gould spends a lot of time debunking the myth of objectivity as a psychological characteristic of scientists. As anyone who studies science soon begins to realize, scientists are not very objective when it comes to their own work, but group objectivity is what matters. Individual scientists may lack objectivity when it comes to their own pet hypotheses, but others will happily take up the slack. Science is organized so that subjectivity can be reduced, resulting in as much objectivity as scientists need.

Gould warns of the great harm done in science and elsewhere by caricaturing one's opponents. The contrast between the traditional ‘positivistic’ views of science and more recent ‘postmodernist’ views is ripe for caricature. Supposedly, positivists think that scientists are infallible and provide absolute truth, whereas postmodernists insist that scientists are driven to come up with the views that they do primarily by such social forces as sexism, racism and homophobia. Gould tries to skate between these two extremes. Science is socially embedded, but the recognition of this fact can only aid scientists in their goal of recording and explaining the natural world.

A second contribution that the humanities can make to science is to help scientists improve their communication skills. Gould thinks that academics in general do not write well, with scientists especially deficient in this respect. One of the things that made Gould stand out from his contemporaries was that in general he wrote as well as any humanist ever has. Young academics in departments of English are stuck teaching courses in remedial writing, just as those of us in departments of philosophy are stuck teaching courses in remedial thinking, but neither group is likely to take much satisfaction from the goal that Gould has assigned them.

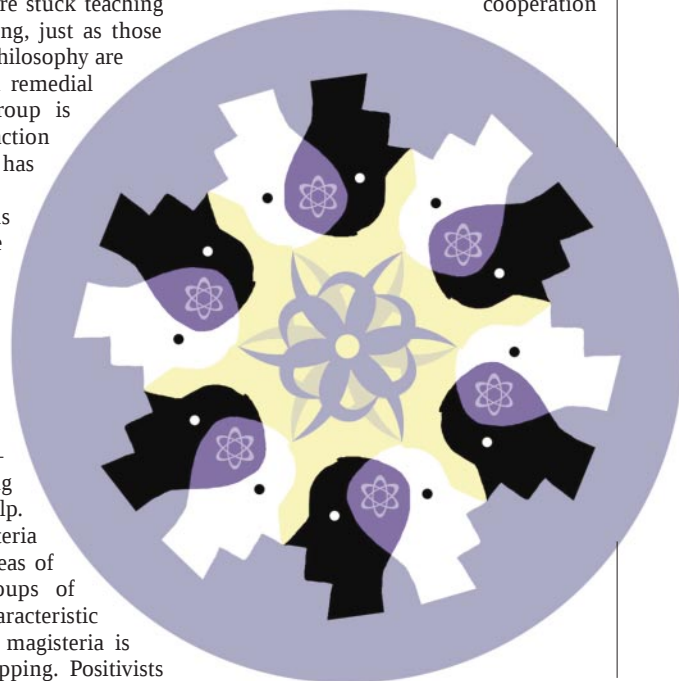
Finally, Gould claims that those of us in the humanities are charged with setting out the proper boundaries of all magisteria, including science. One of the weaknesses of this book is that Gould does not say enough about his central notion — magisteria — and turning to a dictionary does not help. As far as I can tell, magisteria are something akin to areas of expertise shared by groups of experts. About the only characteristic that Gould attributes to magisteria is that they are non-overlapping. Positivists

took it as one of their tasks to define science. Quite a few philosophers of science today think that attempting to draw a line between science and everything else is not very helpful. There is no ‘essence’ to science, no set of attributes that characterizes all scientists and only scientists throughout all time. Science has evolved and continues to evolve. For some reason, Gould thinks that our concepts must be absolutely sharp to be of any use at all. Despite having a history, Gould argues, they have and must have an essence as well.

A second weakness of this book is that it is largely a collection of essays and parts of essays that have been welded into a single narrative. The seams show. Readers might wonder about the quirky, although arresting, title. The contrast between the hedgehog and the fox is initially meant to distinguish between being very good at one thing (the hedgehog curling up in a ball when attacked) and reasonably good at many things (the fox). He refers to this metaphor frequently in his book, but for me it doesn't add much to his exposition, and the Magister's pox hardly merits a mention except to rhyme with fox. Gould expends much effort arguing that science, once it is properly understood, is not in conflict with the humanities. But he says little about the help, if any, that science can give to the humanities.

Another reason why Gould introduces the notion of magisteria is to help usher in the Age of Aquarius, when peace will guide not only the planets but also academic disciplines. He himself has been badly burned, especially in connection with the controversy over E. O. Wilson's ‘sociobiology’. Wouldn't everyone benefit if we all worked together and the lamb were to lie down with the lion?

I too like peace and quiet,
cooperation



and generosity, and all areas of human endeavour can at times be characterized in these terms, but not often.

Why are people so loath to apply to themselves the basic principles that they apply to everyone else? Because it all depends on whose ox is being gored. Throughout this book, Gould preaches peace. In particular, he would like to end his conflict with Wilson, but in the final third of this book, Gould cannot resist taking one last swipe at his long-time adversary. Although Gould is as gentle and gentlemanly as possible, he attacks Wilson's reductionism under the guise of William Whewell's 'consilience of inductions'. A consilience of inductions occurs when the data used to support one hypothesis also turn out to support another hypothesis as well. Gould argues that Wilson has got Whewell all wrong: Whewell's notion of consilience is not in the least reductionistic.

In his final collection of essays, *I Have Landed*, Gould could not help wondering what comes next. Unfortunately, what came next was a second bout of cancer and premature death. But had he beaten the odds a second time, I think Gould would have continued to do battle with Wilson. I picture two warriors sinking into quicksand as they flail away at each other one more time — just one more time. ■

David L. Hull is emeritus professor in the Department of Philosophy, Northwestern University, 1818 Hinman Avenue, Evanston, Illinois 60208-1315, USA.

Einstein's view of Germany

Einstein in Berlin

by Thomas Levenson

Bantam Doubleday Dell: 2003. 400 pp. \$25.95

Freeman J. Dyson

Thomas Levenson is a film-maker who produces documentary films for public television. He has a sharp eye for the dramatic events and personal details that bring history to life. His latest book is a social history of Germany from 1914 to 1933, when Albert Einstein lived in Berlin. The picture of the city's troubles comes into clearer focus when viewed through Einstein's eyes. He was a good witness, observing the life of the city in which he played an active role but from which he remained emotionally detached.

Einstein wrote frequent letters to his old friends in Switzerland and his new friends in Germany, recording events as they happened and describing his hopes and fears. His daily life and activities come intermittently into the narrative, but are not the main theme. The main theme is the tragedy of the First World War, a tragedy that began in 1914 but



did not end with the war in 1918. It continued to torment the citizens of Berlin until 1933, and led them finally to put their fate in the hands of Hitler. Hitler was able to gain power by promising to erase the tragedy and bring back the happy days of the empire when Germany was prosperous and united.

"There are, to be sure, too many biographies of Einstein and not enough of Poincaré," writes Peter Galison in a forthcoming book, *Einstein's Clocks and Poincaré's Maps*. Every aspect of Einstein's life — the personal, the political, the scientific and the philosophical — has been described in detail and analysed in depth by his various biographers. The world does not need another Einstein biography. Fortunately, *Einstein in Berlin* is not a biography. Levenson has borrowed everything he needs from the published correspondence and the existing biographies of Einstein, with full acknowledgements and an excellent bibliography. The new and original aspect of this book is the context in which Einstein is placed: an in-depth study of the social pathology that gripped Berlin from the day that Einstein arrived in 1914 to the day he left in 1932.

The tragedy is a play in two acts: the first act covers the years of war, and the second act the years of the Weimar Republic. The most remarkable feature of the first act is the general belief among Einstein's friends in Berlin that the war was winnable. The war was widely welcomed as an opportunity for Germany to achieve its proper status as a great power. Einstein observed that his academic friends and colleagues were even more deluded with patriotic dreams of grandeur than the ordinary citizens that he met in the street.

For example, in June 1918, after the last great German offensive on the Western front

had failed, Felix Klein, a mathematician second only to David Hilbert in fame and authority, spoke in Göttingen to an audience of leading industrialists and government officials. He talked confidently of the coming victorious conclusion of the war, of the harmonious collaboration of German science with industry and the armed forces, and of the expected increase in support for mathematical education and research after the victory was won.

The state of mind of the mandarins in Berlin was very different from that of their enemies in Paris and London. In Paris the war was seen as a desperate struggle for survival. The guns on the Western front were so close that everyone in Paris could hear them. In Britain the war was seen as a tragedy that had done irreparable harm to Britain and to European civilization, no matter who won it. When the war came to an end in November 1918, the British public looked back on it as an unspeakable horror that should never under any circumstances be allowed to happen again. But a large part of the German public looked back on it differently, as a test of strength that they could have won if they had not been stabbed in the back by traitors at home. This book explains how that fatal German sense of betrayal came into being.

The second act of the tragedy is the story of the slow collapse of the Weimar Republic and the rapid rise of Hitler. Einstein was a firm supporter of the republic, but he saw which way the wind was blowing. One episode in the tragedy epitomizes the whole story. Erich Remarque's book *Im Westen Nichts Neues* was published in 1929 and immediately became an international best-seller. It is the finest fictional account of the First World War, as seen through the eyes of a group of young Germans who die pointlessly

in the carnage of the Western front. In 1930 it was made into a Hollywood film, *All Quiet on the Western Front*, which was shown all over the world — except in Germany. When the film's distributors tried to show it in Berlin, Hitler's friend Joseph Goebbels organized a riot in the cinema. Further Nazi demonstrations and violent protests against the film followed. The Weimar government then banned the film throughout Germany because the Nazis considered it unpatriotic.

This episode explains a mystery in my own family. One of my relatives, who is now 94 years old, has lived in Germany all her life and grew up in the Weimar years. Many years ago I gave her Remarque's book to read and she found it very moving. "This book is wonderful," she said. "Why didn't they let us read it when it was published? That was before the Hitler time, but we were told it was disgusting and shameful, and that respectable people should not read it." So the respectable Germans of her generation, even those who were not Nazis, did not read Remarque. I had always wondered why, and now I know. ■

Freeman J. Dyson is at the Institute for Advanced Study, Princeton, New Jersey 08540, USA.

Getting heated over glaciation

Snowball Earth

by Gabrielle Walker

Crown Publishers: 2003. 256 pp. \$24.95

Bloomsbury: 2003. £16.99

Euan Nisbet

Snowballs can be fun. But when they have rocks in them, the arguments start, just as between Calvin and Susie in the comic strip *Calvin and Hobbes*. *Snowball Earth* is a Hobbes-eye view of the succession of missiles, ambushes and strategies in the long-running skirmishes over the idea that the Earth may have become so cold 600 million years ago that even the tropics froze over. The scene is superbly described, and each throw is meticulously followed from the gathering of the snow to the splat on the target. But a warning: it gets personal and gossipy — snowballologists will need many measures of generous spirit to read this.

The snowball debate first rumbled around the darker recesses of the Spitsbergen room of Cambridge University's Sedgwick Museum 30–40 years ago. Brian Harland had suggested that there had been a great glaciation just before the start of the Cambrian period. Few believed him beyond the confines of the Sedgwick Museum, and even within I recall scepticism (including my own). The new ideas of plate tectonics could surely explain most things, even glacial sediments closely juxtaposed to tropical

deposits. We knew that rapid continental motion occurred: India's march northwards showed that. Nevertheless, Harland serenely held to his views, supported later by Mike Hambrey. Then Joe Kirschvink published the briefest of notes on 'Snowball Earth', lost in a fine book, *The Proterozoic Biosphere*, which was widely read but so vast that its most common use is as a doorstop. There languished the notion.

The Canadian government intervened, deciding in its wisdom to deconstruct what was then arguably the finest (and economically most useful) national geological survey on Earth. Paul Hoffman, a distinguished Arctic field geologist, was unable to work as usual. He departed south and turned his attention to Namibian rocks laid down in the latest Proterozoic, the time just before the Cambrian. What he found were rocks deposited in glacial conditions. But puzzlingly, these rocks were closely juxtaposed with carbonate sediments laid down in warm conditions. Moreover, the succession was very similar to the Spitsbergen rocks that Harland's group had studied.

The 'snowball' idea states that around 600 million years ago there were episodes when even in the tropics the ocean surface froze. The white, icy planet then reflected so much sunlight that the condition persisted. Life luckily survived, presumably in Noah's Arks of habitat in warm water around hot springs or photosynthesizing in open channels in the icy ocean. Eventually, as carbon dioxide continued to degas from volcanoes, the atmospheric greenhouse became so strong that tropical melting began. As ice melted, the newly darkening planet would have absorbed heat, a positive feedback leading to runaway warming. This would have been exacerbated by the release of dissolved carbon dioxide from the heating oceans and methane from vast hydrate stores, so the snowball suddenly flipped to a hothouse hell.

Gabrielle Walker recounts in racy detail the chain of discovery. Hidden in the social froth, she provides nuggets of gold: simple explanations of complex evidence. Carbon isotopes, facies analysis, methane hydrates, metazoan evolution — complicated topics all — slip down with spoonfuls of scandal sugar. The book lucidly exposes the observations that shaped the hypothesis and its testing, the critics and the bases of their criticisms, and the subsequent

growth of the idea as the debate provoked more and more insights. Just as Sir Harold Jeffreys' opposition to continental drift was crucial in provoking a rigorous theory of plate tectonics, so Walker shows clearly how disputes about the snowball (for example, the contrary views of Nick Christie-Blick), brought deeper understanding.

Did Snowball Earth really happen? Sitting next to a scientific knight of excellent judgement at the end of Hoffman's peroration on receiving the Wegener Medal in 2001, I noted his warm applause (did he stamp his sandals? Was he clapping act or fact?), while I was still weighing the balance. Snowball Earth is entering the textbooks. Walker quotes a wildly improbable (but splendidly diverting) speculation that Hoffman will become a lord. In reality, like most scientists, geologists welcome but do not seek applause or rewards (or they'd have chosen easier and more remunerative jobs): their thrill is the chase. The details (snowball or slushball) remain hotly argued, and there are many false alarms and excursions, but there is an emerging consensus that something odd is being hunted down.

Is it important? The notion that Earth's climate has alternative stable states and can flip between them carries a warning for today. The stable climate of the past 10,000 years may be a freak — the hidden nightmare monsters that made the snowball may still be out there, especially methane. The Arctic sea floor and the continental shelves worldwide have widespread methane hydrates that will be released by warming. We are already turning Spitsbergen into fried ice-cream. Is it time to learn from geology's nightmares? Kyoto began the estrangement between the United States and Europe: maybe it could also begin the reconciliation. To get technical, either by a stand-alone methane agreement, or by resetting Kyoto's timescale for calculating global-warming potential from 100 years to a 'poor-nation-

friendly' 20 years, and hence by placing more emphasis on vigorous cuts in methane emission,



we may find a cheap-to-implement accord, acceptable to the United States, the European Union and poorer nations, that actually tackles radiative warming now, within a few decades.

Noah's rainbow was a covenant for all living species, not just humanity. Humanity's emissions now wilfully block the infrared part of that promise. If our nations corrupt the rainbow, is it time to search out supplies of gopher wood and start collecting eukaryotes two-by-two, trying to keep our genetic company afloat while the greedy world goes into liquidation? In the meantime, *Snowball Earth* makes good spring reading. And when you are done, pass it on to any handy teenager, who will be enthralled, if not by the science, then by the storytelling. ■

Euan Nisbet is in the Atmospheric Group, Department of Geology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.

Gould and God

A Devil's Chaplain

by Richard Dawkins,
edited by Latha Menon

Weidenfeld & Nicolson: 2003. 320 pp.
£16.99

Jerry A. Coyne

The original 'Devil's chaplain' was Robert Taylor, an apostate priest and self-styled 'infidel missionary' whose rabble-rousing entourage stormed into Cambridge in 1829 — halfway through Charles Darwin's undergraduate career. Taylor's brew of atheism and republicanism held no allure for young Darwin, then a conventional Christian thoroughly steeped in the hierarchical values of nineteenth-century Britain. But Darwin never forgot Taylor's nickname. In 1856, pondering the brutal inefficiency of natural selection, Darwin wrote to his friend Joseph Hooker: "What a book a devil's chaplain might write on the clumsy, wasteful, blundering, low and horribly cruel works of nature!" Although creationists constantly remind us that *On the Origin of Species* is a satanic work, Darwin never aspired to succeed Taylor: he was too shy of controversy, too worried about the happiness of his devout wife.

Richard Dawkins, however, is more than happy to step into Taylor's shoes. Religion, writes Dawkins, is a "malignant infection" of the human mind. Six of the 32 essays in this eclectic collection (culled largely from the British press) address religion either directly or indirectly, a surprising statistic for a science writer whose chair at Oxford University is dedicated to the public understanding of science. However, another six essays attack further impassable routes to knowledge

(such as homoeopathy, crystal worship and postmodernism). Clearly, Dawkins sees his brief as not only popularizing science, but demolishing its competitors.

Dawkins' other major bugbear — and the main non-divine theme of *A Devil's Chaplain* — is the late Stephen Jay Gould. Dawkins offers here a gracious tribute to Gould, but also turns his incisive analytical mind and lucid prose against Gould's ideas with devastating effect. For example, Gould's notion that the evolutionary processes that shaped the early history of life (with new phyla being produced) differ qualitatively from the processes that direct more recent evolution (with only lower-level taxa appearing) is deflated in two deft sentences: "It is though a gardener looked at an old oak tree and remarked wonderingly: 'Isn't it strange that no major boughs have appeared on this tree recently. These days, all the new growth appears to be at the twig level!'"

In fact, the critiques of Gould and God are not as unrelated as they appear. As this collection makes clear, Dawkins is a fierce advocate of scientism, the philosophy that genuine truths — as opposed to spiritual or personal 'truths' that are not universally held — can be found only through the scientific method. Gould was not as strict: he was an accommodationist, insisting in works such as *Rocks of Ages* that both religion and science are independent and valid domains of enquiry. Dawkins, then, is both scientifically and philosophically opposed to Gould. Several books have been written about the scientific arguments between Gould and Dawkins, but Dawkins aficionados outside Britain have had little exposure to his withering assaults on religion, pseudoscience and accommodationism. It's a rare treat to see him sail into battle, prose and logic perfectly attuned to the destructive business at hand. I should add that *A Devil's Chaplain* is not wholly focused on religion and Gould: it also includes obituaries of friends, and essays on ethics, genetic determinism, Africa (Dawkins spent much of his childhood in Kenya) and meme theory (the collection's only low point, given my view that 'mimetics' is an extended tautology that has yielded no real understanding of human culture). But it's in religious territory that *A Devil's Chaplain* is most compelling.

"Modern theists," writes Dawkins, "might acknowledge that, when it comes to



Baal and the Golden Calf, Thor and Wotan, Poseidon and Apollo, Mithras and Ammon Ra, they are actually atheists. We are all atheists about most of the gods that humanity has ever believed in. Some of us just go one god further." But Dawkins goes beyond a mere defence of atheism. He also subscribes to the American writer H. L. Mencken's dictum that: "We must respect the other fellow's religion, but only in the sense and to the extent that we respect his theory that his wife is beautiful and his children smart." Why, asks Dawkins, should the public give religious arguments any more credibility than arguments for other brands of non-scientific 'truth'? Curiously, Dawkins does not explore why religious ideas get undue respect. Surely one reason is that arguing about religion (especially when one participant is an atheist) is unproductive, likely to produce only mutual dislike. No *rapprochement* is possible between those whose beliefs derive from evidence and those whose beliefs either do not depend on evidence or are unshaken by contrary evidence. This is why science and religion are incompatible ways of viewing the world.

Dawkins' critique of religion rests on three points. First, because different faiths make very different claims about the world, they cannot all be true; and none of the claims (such as the bodily assumption of Mary into heaven) can be scientifically verified. Second, the choice among faiths is not based on rational consideration: the vast majority of people simply practice the religion of their parents. This is especially galling to Dawkins, who sees the easy indoctrination of children as a product of natural selection favouring the rapid spread of information between generations. Finally, Dawkins considers religions to be vehicles of evil because they facilitate the labelling

of people as either 'us' or 'them', fostering xenophobia and its attendant horrors — Northern Ireland and the Middle East come to mind.

These views are summarized in a wonderfully passionate essay, "Time To Stand Up", written shortly after 11 September, 2001. One excerpt: "To label people as death-deserving enemies because of disagreements about real-world politics is bad enough. To do the same for disagreements about a delusional world inhabited by archangels, demons, and imaginary friends is ludicrously tragic."

Would that there were an afterlife, so that Robert Taylor could smile upon his far more effective heir! As Taylor and his fellow freethinkers knew, atheism in early nineteenth-century Britain was blasphemy and thus illegal: Taylor was twice jailed for his activities. Thankfully, such strictures are now much rarer, but a subtler form of repression prevails in places such as the United States. Scientist-atheists, bowing to prevalent notions of politically correct social inclusiveness, are unwilling to express their opinions for fear of offending religious sensibilities. But Dawkins makes a strong case that most religions are insidious and dangerous illusions. It's time for those who agree to stand up beside him. ■

Jerry A. Coyne is in the Department of Ecology and Evolution, University of Chicago, 1101 E. 57 Street, Chicago, Illinois 60637, USA.

The writing on the slate

Nature via Nurture: Genes, Experience and What Makes Us Human

by Matt Ridley

Fourth Estate: 2003. 320 pp. £18.99

HarperCollins: 2003. \$25.95

Andrew Berry

Is who we are determined ineluctably by our biological inheritance or, more malleably, by our experience? The debate is surely as old as human consciousness. In 1874 Francis Galton gave it its modern identity when, borrowing from Shakespeare's villain Caliban, "a devil, a born devil, on whose nature nurture can never stick", he cast the issue in terms of what he called a "convenient jingle of words": nature and nurture.

Having made the distinction, in *Hereditary Genius* Galton then set the tone for the debate to come by hewing dogmatically to an extreme position: "I have no patience with the hypothesis occasionally expressed, and often implied, especially in tales written to teach children to be good, that babies are born pretty much alike, and that the sole agencies in creating differences between

boy and boy, and man and man, are steady application and moral effort. It is in the most unqualified manner that I object to pretensions of natural equality."

The other extreme has also attracted its own inflexible adherents, most notably members of the 'behaviourist' school founded by J. B. Watson, whom Ridley quotes: "Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select: doctor, lawyer, artist, merchant-chief, and yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors."

This polarization remains with us to this day: the debate is typically couched in terms of nature versus nurture, implying that these factors are mutually exclusive. The issue is clouded by the difficulty of bringing conclusive evidence to bear. Experiments on humans are impracticable or unethical, although the Moghul emperor Akbar, unfettered by regulations on the use of human experimental subjects, did apparently raise several individuals in total isolation to determine which religion — Hinduism, Islam or Christianity — they would spontaneously embrace. The experiment was inconclusive: the lack of stimulation during their development turned Akbar's unfortunate human guinea-pigs into deaf mutes.

Modern commentators have struggled to extricate themselves from the straitjacket of Galton's dichotomy. Some genes do indeed act independently of the environment: regardless of my lifestyle or where I live, I will inevitably develop Huntington's disease if I carry the disease-causing mutation. And conversely, plenty of our behaviour is largely environmentally determined — that I speak English, not Turkish, is simply a reflection of where I was raised and by whom. But not all behaviour resides at one or other end of the spectrum: genes and the environment often interact such that the either/or categorization of the 'versus' view is misleading.

However, as the evolutionary biologist David Sloan Wilson pointed out in the *New York Times* on 25 February 2003, the rhetorical allure of the extremes remains strong: "Everyone calls themselves an interactionist. Yet often, when you

scratch below the surface, you find a sociobiologist who marginalizes the importance of culture, or a social constructivist who hates the very idea of sociobiology, and they end up painting caricatures of each other. True integrative thinking is in the very early stages."

Nature via Nurture is a book-length exercise in 'integrative thinking': science writer Matt Ridley has produced a paean to interaction that will do much to erode the mutually exclusive view of nature and nurture.

Interaction is best exemplified in a simple idea that typically makes an appearance somewhere near the beginning of a genetics textbook and is then ignored throughout the rest of the book: the outcome produced by a gene may depend upon the context in which the gene is expressed. Citing new work by Darlene Francis at Emory University in Atlanta, Ridley provides an extraordinary and elegant example. C57 and BALB strains of mice differ discretely in some aspects of adult behaviour. But C57 embryos transplanted to BALB uteri and raised by BALB mothers display, as adults, aspects of BALB behaviour; mere cross-fostering (C57 to BALB parent) after birth, however, does not provoke the change, implying that the uterine environment is the critical context. The

C57 genotype expresses C57-typical



behaviour only after development in a C57 uterus.

Ridley's historical tour of several disciplines is a delight: the pivotal players in ethology, neurobiology, anthropology and psychology are brought to life in engaging pen-portraits. We encounter anthropologist Franz Boas in 1884 during his first field season among the Inuit of Baffin Island as he notes in his diary: "These are the 'savages' whose lives are supposed to be worth nothing compared with a civilized European. I do not believe that we, if living under the same conditions, would be so willing to work or be so cheerful and happy." Konrad Lorenz appears both in his best-remembered guise, pursued by a string of adoring ducklings, and in an altogether more sinister one. While working as a military psychologist in Poland in 1942, Lorenz participated in SS-sponsored research designed to distinguish between inferior Polish and superior German characteristics in 'half-breeds'. Nor does J. B. Watson of behaviourism fame fare too well in retrospect: his fall from grace was occasioned by an extramarital love affair, and he ended up, appropriately enough, applying his skills in pavlovian conditioning to advertising Johnson's baby powder. The human detail enriches *Nature via Nurture*, but Ridley by no means subscribes to the modern dumb-it-down school of science writing, in which the science itself becomes a sideshow to the serious business of prying into the scientists' personal lives.

Ridley's book reminds us of the importance of good science writing. Because he is not a professional scientist, Ridley is not stuck deep in a disciplinary trench and has the freedom to range across huge swathes of intellectual territory. In doing so, he has given us a rich overview and a compellingly integrated picture of a great deal of science, both old and new. Make *Nature via Nurture* part of your nurture. ■

Andrew Berry is at the Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA.

The evolutionary blackbird

The Story of Life

by Richard Southwood

Oxford University Press: 2003. 272 pp. £19.99

Carl Zimmer

The poet Wallace Stevens wrote about 13 ways of looking at a blackbird. Perhaps someday another poet will write about 13 ways of looking at the history of life. I can certainly think of 13 different scientists who have written books on the subject, each of which is coloured by its author's expertise.



Books by vertebrate palaeontologists are dominated by animals with bones, despite the fact that vertebrates make up a tiny proportion of the world's biodiversity today — not to mention the fact that they didn't exist for the first 3 billion years or so of life's history. The Precambrian expert Andrew Knoll has looked at those first 3 billion years in great detail in his new book *Life on a Young Planet* (to be reviewed in *Nature* shortly), leaving the details of dinosaurs and mastodons to others. For yet another take, try *The Major Transitions in Evolution* by John Maynard Smith and Eörs Szathmáry. For them, the essence of life's history is the emergence of new kinds of complexity. Instead of fossils, you get equations.

In *The Story of Life*, Oxford ecologist Richard Southwood takes his own look at the evolutionary blackbird. Southwood is a leading figure in ecology, thanks to his seminal work on insects and his landmark book *Ecological Methods*. For 18 years he taught an introductory course on the history of life, and out of that experience has sprung *The Story of Life*. Not surprisingly, Southwood sees the history of life in an ecological light, not as a single-file parade of new life forms but as a network of species whose links are being perpetually reworked.

This network first took shape over 3 billion years ago, as early microbes cooperated to harness the energy in their environment. The network grew more complex as animals and other multicellular organisms evolved, and as reefs offered new ecospace for species to colonize. Southwood recounts how dry land was transformed over hundreds of millions of years, as bacterial crusts gave way to forests that offered a new ecospace as vast as that of coral reefs. Over time, ecosystems change like gently tapped kaleidoscopes, Southwood writes, although mass extinctions give them a good shake from time to time.

Southwood displays an impressive sweep

of knowledge about life, from the fauna of hydrothermal vents to the anatomy of plant-eating birds' digestive tracts. For the most part, he has kept abreast of the latest developments in evolutionary research, although from time to time he slips back into comfortable textbook explanations. Describing the great domains of life, for example, he writes: "The Archaeobacteria also fall into two groups, both of which have lifestyles that are very unusual, but which could have been maintained on the ancient earth." These microbes (which are now generally called Archaea, not Archaeobacteria) can indeed be found in strange places, such as geysers and oxygen-free swamps. But they can also be found in ordinary places, such as grassland soil and the open ocean, where they outnumber bacteria. Archaea got a reputation for being bizarre only because scientists discovered their more unusual members first.

A more serious shortcoming in *The Story of Life* is the scant attention paid to DNA, which has revolutionized our understanding of evolution's course. The ecological changes that Southwood details were made possible by changes to genes, and scientists are starting to get some hints of what those changes were, from the promiscuous gene swapping between early microbes to the recruitment of old genes to make new structures such as jaws and fingers.

Despite these grumbles, I recommend *The Story of Life* to those looking for a swift, efficient delivery of the most important information we have on how life has blossomed on Earth. Southwood is succinct and clear, and his narrative rarely gets bogged down with historical digressions or personal anecdotes. Although this style has its strengths, it also has its weaknesses. At the beginning of his book, Southwood claims that the story of life "provides a benchmark for judgments on the environmental problems of today". But when Southwood finally reaches our own age, he seems almost

indifferent as he zips through the ways that we are altering the climate and the planet's ecosystems. A little more passion would have been welcome. After all, these days the evolutionary blackbird is beginning to look more like a canary in a coal mine. ■

Carl Zimmer is a freelance writer who lives in New York. His most recent book is *Evolution: The Triumph of an Idea*.

Metabolic gardening

Wandering in the Gardens of the Mind: Peter Mitchell and the Making of Glynn

by John Prebble and Bruce Weber
Oxford University Press: 2003. 344 pp.
\$65, £45.50

E. C. Slater

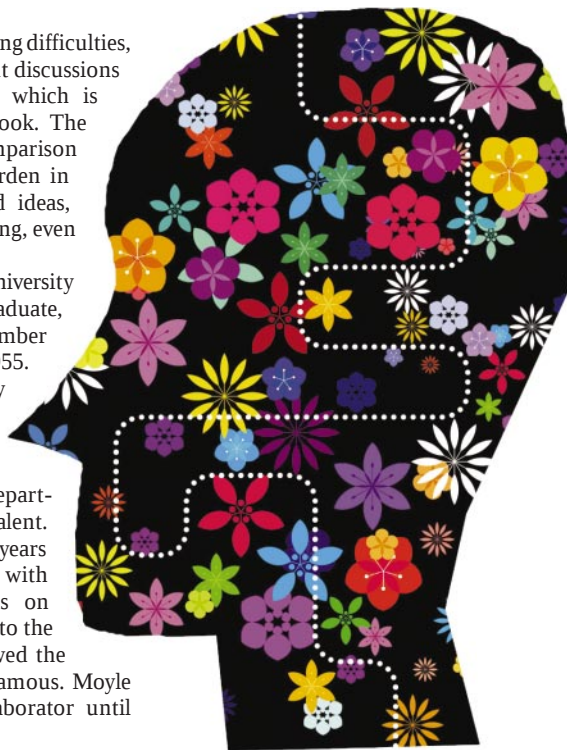
Peter Mitchell was an outstanding theoretical biologist who had what Leslie Orgel has characterized as a counterintuitive idea that led to a fundamental shift in the way that biochemists viewed energy metabolism. The story of the gradual replacement of the chemical hypothesis of energy transduction, which involved a non-phosphorylated high-energy intermediate (not a phosphate compound, as stated in this book), with the chemiosmotic theory has become, for historians of science, a classic example of how a paradigm shift occurs.

In this biography, the story is related again, primarily from Mitchell's viewpoint.

Because of ill-health and hearing difficulties, Mitchell preferred to carry out discussions by lengthy correspondence, which is extensively quoted in this book. The title refers to Mitchell's comparison of the human mind to a garden in which are planted facts and ideas, which one keeps on rearranging, even when asleep.

Mitchell worked in the University of Cambridge as an undergraduate, research student and staff member for 16 years from 1939 to 1955. He did not have a particularly distinguished academic record, nor does he seem to have been completely at home in the biochemistry department, despite its wealth of talent. However, in his last few years there he published, together with Jennifer Moyle, experiments on the transport of phosphate into the bacterial cell that foreshadowed the work for which he became famous. Moyle remained his scientific collaborator until her retirement in 1983.

There are some errors in the description of the Cambridge environment concerning the Molteno Institute of Parasitology and Biology, headed by David Keilin, where Mitchell is described as having found "shelter" (possibly in late 1945 to early 1946). E. F. Hartree was not a member of the technical staff, as stated in this book, but a PhD from the University of London and co-author of most of Keilin's classic papers for over 20 years. I am described as an "occasional post-doctoral researcher" in the institute, which



is surprising, considering that I worked there from late 1946 to 1955, during which time I published the chemical hypothesis of oxidative phosphorylation in *Nature* in 1953. It is the overthrow of this hypothesis by Mitchell that is a major theme of this book.

In 1955, Mitchell moved to the University of Edinburgh, where he developed the first version of the chemiosmotic theory, which was published in *Nature* in 1961. Illness soon caused him to move to Cornwall, where he converted a derelict manor house called Glynn House into a private laboratory as "a quiet haven for untrammelled scientific work and thought", as he wrote to me in 1964. The story of Glynn is the second strand in this biography. In 1966, Mitchell published, in what became known as 'The Grey Book', a modified version of the chemiosmotic theory in which he introduced the concept of proton motive force and 'loops' in the electron-transfer chain.

It took about ten years for the chemiosmotic theory to be generally accepted, and in 1978 Mitchell received the Nobel Prize in Chemistry. Mitchell was so confident that oxidative phosphorylation was solved that in 1976 he resolved to cease biochemical research and move into a new field, perhaps economics. The following year the decision was reversed when he considered that his theory was again under attack. It is worth noting that when the chemiosmotic theory became the new paradigm, Mitchell was just as resistant to attacks on parts of the theory as the supporters of the old paradigm had been.

The first assault came from Mårten

History in the making

It was one of the most momentous moments in the history of science, but even those people involved didn't know for sure just what its impact would be. *50 Years of DNA*, published by Palgrave Macmillan this month, tells the story of how the structure of DNA was discovered, and describes its subsequent impact on science, medicine and culture.

The book includes facsimiles of the landmark paper by James Watson and Francis Crick (pictured far left), along with those by Maurice Wilkins and Rosalind Franklin (pictured centre) and their colleagues that first appeared in *Nature* on 25 April 1953.

Introductory chapters take the reader on a journey from the early days of molecular biology through to the sequencing of the human genome, incorporating interviews with some of the key players. The book concludes with a collection of essays by prominent experts giving



their views on the history and significance of Watson and Crick's discovery, along with their predictions of what genetics might achieve in the next 50 years. Edited by Julie Clayton and Carina Dennis, both former editors at *Nature*, the book celebrates the triumph of the DNA double helix.

Wikström's experiments showing proton pumping by cytochrome oxidase, published early in 1977. But it was not until 1985 that Mitchell was prepared to accept the evidence presented by Wikström and others, and then only when he had thought up a mechanism (since shown to be incorrect) to accommodate it. In Mitchell's garden of the mind, ideas seem to have deeper roots than facts.

The second attack concerned the chemical mechanism of ATP synthesis. Mitchell's postulate (from which he never wavered) that protons are directly involved in the esterification of ADP by phosphate was challenged by Paul Boyer, Harvey Penefsky and my group, who believed that ATP synthesis is brought about by a conformational change in the ATPase such that ATP, already formed spontaneously on the enzyme, is released. This view led to the Nobel Prize in 1997 for Boyer and John Walker.

Have we reached the end of the story? Not quite. The postulate of a non-phosphorylated high-energy intermediate in the chemical hypothesis is correct; its nature is not. The postulate of the chemiosmotic hypothesis that this intermediate is an electrochemical proton gradient is correct; the way in which it was thought to synthesize ATP is not. The binding-change mechanism of ATP synthesis proposed by Boyer is probably correct; but will the rotary mechanism depicted in the textbooks survive challenges?

I recommend this biography as a scholarly account of the life of one of the giants of late-twentieth-century biochemistry who was also a fascinating, if enigmatic, human being. His amiable character, his eccentricities and his health and emotional problems are fully described. But I still do not think that I really understand the workings of his mind.

Like Tom Blundell, writing in the foreword to this book, I find it quite extraordinary that Mitchell refused to permit Bob Williams to publish their correspondence from before the publication of Mitchell's first paper on the chemiosmotic theory, concerning the possible role of protons in ATP synthesis. The reason given for Mitchell's 1977 reversal of the decision to cease biochemical research — namely that he believed “there was a strong movement among the antagonists of the general chemiosmotic theory... to take advantage of our withdrawal and press new, ill-founded research claims, the effect of which was to undermine the consensus of opinion which we had been working on to promote” — I find even more extraordinary.

But perhaps it is not possible for a mere biochemist to analyse the mind of a genius. ■
E. C. Slater is emeritus professor of physiological chemistry at the University of Amsterdam, the Netherlands. He is currently at 9 Oaklands, Lymington SO41 3TH, UK.

Science in culture

Art after DNA

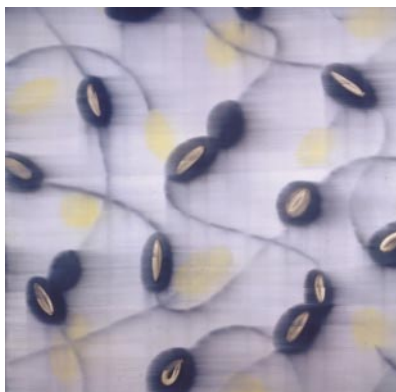
The double helix has inspired scientists and artists alike.

Lynn Gamwell

No sooner had James Watson and Francis Crick published their seminal paper on DNA in 1953 than artists began depicting the double helix as a cultural icon. Salvador Dalí painted swarms of spiralling DNA molecules in his paintings from the late 1950s, such as *Butterfly Landscape (The Great Masturbator in a Surrealist Landscape with DNA, 1957–58)*, and decades later sculptor Tom Otterness forged a bronze double helix joined by tiny stick figures (*DNA Chain, 1986*). Some artists responded to some of the troubling issues surrounding DNA, such as cloning and stem-cell research. Alexis Rockman, for example, painted a colourful landscape populated with genetically modified plants and animals to warn viewers that the future may contain mutated monsters (*The Farm, 2000*).

Although art about science icons and applied science is interesting, it does not focus on the pure science of DNA. That topic is addressed by artists who express wonder at the highly complex, ever-changing organic processes that are the physical mechanism for the force of life. One example is the British painter Mark Francis, whose *Source* (1992) is a wall-size depiction of sperm cells presented not with clinical accuracy but out of focus, as if seen through a veil. Inspired by microscopy, Francis uses a vocabulary of curved, flat shapes that he inherited from nineteenth-century art nouveau designers — who, in turn, first copied the shapes from stained, transparent slices of tissue prepared between glass plates for viewing with a light microscope. Francis uses a modern electron microscope for his source images, but his forms retain the flat, free-form appearance associated with a century of biomorphic abstract art. In a nod to recent methods of charting complex genomic data, Francis paints the sperm cells hovering above a grid.

In the 50 years since Watson and Crick cracked the structure of DNA, biologists and bioinformaticists, have painstakingly mapped and sequenced the genomes of various species: the fruitfly, the mouse



Sperm cells: the view of sperm in *Source* by Mark Francis was inspired largely by microscopy.



Genetic screen: Benjamin Fry's *Valence* uses the BLAST algorithm to create moving patterns.



DNA 2 by Susan Rankaitis is both a stand-alone artwork and a backdrop to the dance of life.

and, in 2001, the human. The computing technology developed for this effort inspired a new generation of video artists, such as Benjamin Fry, a programmer at the Massachusetts Institute of Technology's Media Laboratory, to create visualizations of complex biological data. In *Valence* (2000), Fry has used the BLAST algorithm — a tool for searching through genomes — to produce ever-changing patterns that move rhythmically across a video screen.

The curved free-standing panel *DNA 2*, which measures over 5 metres across, is the result of a collaboration between visual artist Susan Rankaitis, molecular biologist Robert Sinsheimer and choreographer John Pennington. Inspired by the Human Genome Project and tutored by Sinsheimer, Rankaitis combined pictures of DNA with text by the biologist to create her chromosome-shaped collage. It can be viewed alone or used as a backdrop for Pennington, who, attired in coils of pulsing fibre-optic cable, spins across a stage. By projecting flickering light onto her giant chromosome, Rankaitis makes her layered images and text join the dance of life.

Charles Darwin was confident that there was a physical mechanism underlying natural selection, but never found it. Today's artists understand the central role of DNA in this mechanism. It inspires them by embodying Darwin's core idea: that nature is a web of dynamic forces with no predetermined purpose or meaning. Works by artists such as Francis, Fry and Rankaitis resonate with this concept of life, and with the complex, abstract processes that go on silently, systematically and invisibly within the double helix. Lynn Gamwell is director of the Art Museum at the State University of New York, and author of *Exploring the Invisible: Art, Science and the Spiritual* (Princeton University Press, 2002).

B. FRY

S. MCCLAIN/ROBERT MANN GALLERY

TATE PICTURE LIBRARY

Seeking connections

Summarize yourself in the form of a title of a paper in Nature.

Uncovering information in fluctuating systems and understanding unifying principles linking fluctuation phenomena in different systems.

What was your first experiment as a child?

Learning the effect of connecting myself (and my little brother) to electrical sockets in our home, which are 110 volts in the United States, and not so very dangerous.

Who has been the most important mentor in your career?

My first graduate student adviser, Max Delbrück.

What single scientific paper or talk changed your career path?

The private talk Max Delbrück gave to me the day I arrived in his lab in June 1962.

What makes a good scientific mentor?

Teach, by example, how to identify significant scientific problems that are worth addressing. We only live once, so cannot afford time to follow up every interesting idea. We must learn to select problems with the highest probability of a big scientific pay-off.

What gives you the most job satisfaction now? What are your major frustrations?

Satisfaction lies in the search for significant scientific problems. The frustration is the difficulty of making progress with these problems, and the failure of *Nature* to recognize their importance and significance...

Given your interest in the behaviour of economic phenomena, do you have any hot tips for getting rich quick?

Pay attention to the cross-correlations between the movements of different stocks in your portfolio, and update this information regularly. This activity is not unlike those of statistical physicists, where we must pay great attention to statistical correlations between variables, and how these evolve over time.

What is your favourite conference destination, and why?

The Enrico Fermi Summer School at Varenna sul Lago di Como in Italy, because of the emphasis on interactions between conference attendees, both during and after lectures (at meals, at parties ...).

What was the worst/most memorable comment you ever received from a referee?

One referee dismissed a paper as "an incom-

plete derivation of an unimportant result". It turned out to be one of my most important papers (Spherical model as the limit of infinite spin dimensionality *Phys. Rev.* **176**, 718–722; 1968), since cited in many texts on statistical mechanics.

You have the audience in your hands, but some smart-alec asks you the killer question you have no idea how to answer. What's your response?

I would simply admit that I had no idea how to answer the question, throw it open to the audience and invite participants to discuss it later on. After all, the right question can lead to something new.

What book currently resides on your bedside table?

None. I do not read books. Especially not in bed.

What music heads the playlist in your car or laboratory?

Brazilian music, especially from stars such as Daniela Mercury and Maria Bethania.

Assuming the dead can be raised and/or time travel exists, who from the world outside science would you most like to have dinner with?

Jacques (né Jakob) Offenbach, composer of *La Vie Parisienne*, because he loved to laugh and make fun of society.

You are on a plane behind two students obviously going to the same conference, who start to talk about your work. What do you do?

Listen and take notes so that I can remember their perspective: it's a pity that we have no way to know the honest perspective of others on our work. And after they finish, I would identify myself.

The Internet is the bane of scientists' lives because...

The Internet is the opposite of a bane. It is especially vital for interdisciplinary research because it allows us to learn things about other fields without going through the hierarchical authority tree. It also levels the playing field, treating all ages and countries more or less equally.

What would you have become, if not a scientist?

A car thief. I like the idea of doing something that is a challenge, and stealing a car is not unlike stealing secrets from God — that is, from the vast body of scientific law that we do not yet know about. First you must select the car, which requires taste and judgement, weighing up the desirability of the car against



H. Eugene Stanley

Gene Stanley is university professor and professor of physics at Boston University. He likes trying to discover new things about the complex systems in nature, economics and society.

the chance that you can steal it without getting caught (nicer cars have alarm systems, and so on).

What discovery, invention or innovation would most improve your life?

A pill to make me kinder to those I interact with; another to allow me to focus more closely on a single problem; and a third to enable me to see the solution to the problem.

Name one extravagance you can now get away with because of your eminence.

I am sorry to say that physics has no extravagances. I guess the only thing we get away with is not being told we are jerks to our faces, but this is a pity because we lose valuable feedback (so that we have to rely on sitting on planes behind graduate students who are unaware of our identities).

What music would you have played at your funeral?

Romantic soul music, the Black American music of the 1960s.

What's just around the corner?

Better ways of uncovering the useful, 'hidden' scientific information that seems to be carried in fluctuations; to see the connections between the different fields in which fluctuation phenomena are important, such as my recent work on connections between fluctuations in economic systems and fluctuations in systems near special 'critical points' (V. Plerou, P. Gopikishnan and H. E. Stanley *Nature* **421**, 130; 2003). ■

Revelations from a bread mould

Jonathan Arnold and Nelson Hilton

The filamentous fungus *Neurospora crassa* is an organism much loved by geneticists. Its genome sequence has now been unveiled, and includes some surprises — such as the relatively high number of genes.

The lowly bread mould *Neurospora crassa* has a venerable history as a subject of scientific research¹. Its fame really began more than 60 years ago, with work by George Wells Beadle and Edward Lawrie Tatum² that would later win them a Nobel prize. Before then, geneticists were studying genes as hypothetical entities that are transmitted from parents to offspring and — in an unknown way — control inherited traits (as illustrated in Fig. 1)³. Meanwhile, biochemists were looking at proteins and their functions in the cell. There seemed at the time to be little connection between the two disciplines. Working with *Neurospora*, however, Beadle and Tatum made the stunning discovery that genes as units of heredity also encode the proteins that carry out much of the work of the cell.

Now, in a paper that promises to be just as revelatory, Galagan *et al.*⁴ present us with the entire list of genes found in the *Neurospora* genome (page 859 of this issue). And on page 893, Selker *et al.*⁵ take a closer look at a particular type of DNA modification — methylation — in *Neurospora*. They propose that this functions largely to defend the genome against mobile genetic elements.

Galagan *et al.*⁴ report the sequence of more than 38.6 million of the 'letters' — base pairs — from the genetic text of this filamentous fungus. They used the 'whole-genome-shotgun' strategy to read the sequence. The approach involved shattering the genome into more than 200,000 fragments, sequencing the pieces, and then looking for overlaps between the text of the fragments so as to put them together in the right order. This strategy is becoming more and more common, although there has yet to be a convincing demonstration that it produces a puzzle that can be put together correctly. A 'physical map' of landmarks in the *Neurospora* genome is in progress, however; if it tallies with the sequence now revealed, it will provide one of the few convincing tests of the approach.

As is usual at this stage of a sequencing project, there are still a few holes in the sequence (*Neurospora* has been estimated to have a genome of 42.9 million base pairs⁶). But at least 90% of it has been completed. Although the authors have only just finished assembling the text, and so have had little time to meditate on its message, they have already revealed a wealth of new information.

One remarkable finding is the estimated



Figure 1 Following genes from generation to generation in *Neurospora crassa*. The segregation of normal and mutant forms of a gene is shown, with black spores bearing the normal gene and grey spores the mutant form. Each line of eight spores is the product of one cross between two parents. The pattern shows the order in which the spores were formed, and hence the pattern of segregation of the gene (that is, its mendelian inheritance). This cannot be observed directly in most other organisms, explaining why *Neurospora* is the geneticists' favourite child. (Reproduced from ref. 3.)

number of 'words' — protein-encoding genes. Identifying genes in amongst other sequence information is never an easy matter, because there are no accurate procedures for doing this⁷; in the case of *Neurospora*, the quest was particularly hard because more than 50% of the genes are made up of non-contiguous sections, making it difficult to determine where gene boundaries lie. So the gene number may need to be revised in the future. But it seems that, although the *Neurospora* genome is respectively about 4 and 100 times shorter than those of fruitflies and humans, the number of genes is not so different. *Neurospora* has around 10,000 genes, compared with some 14,000 in fruitflies⁸ and 21,000–39,000 in humans^{9,10}. In this sense, we are truly not that far in genetic complexity from the common bread mould.

The function of many of *Neurospora*'s genes remains unknown, however. Comparisons with known genes in public databases are one way of determining function, but it seems that only about 1,400 of the *Neurospora* genes have counterparts in fruitflies, worms and other complex eukaryotes (loosely speaking, those organisms, including humans, whose cells have nuclei). Moreover,

more than half of the *Neurospora* genes have no detectable similarity to those of its yeast relatives *Saccharomyces cerevisiae*¹¹ and *Schizosaccharomyces pombe*, the genomes of which have also been sequenced. Much of *Neurospora*'s genetic text awaits exegesis.

What else does the genome sequence tell us? This fungus shares some significant physical characteristics with other complex eukaryotes, notably a biological clock¹². It knows how to tell the time of day, and much of what we understand about this process in other organisms comes from studies of *Neurospora*. The genome sequence should reveal how the biological clock connects with other cellular processes, such as metabolism and light sensing. Furthermore, the sequence provides hints that, like some plants, *Neurospora* can sense both blue and red light. It was already known that it responds to blue light, and Galagan *et al.*⁴ have now found some additional genes that may be involved in this process. But the discovery of genes that may be needed in sensing red light comes as a surprise. Also, the sequence shows that *Neurospora* shares with yeast several of the core signalling pathways for sensing other aspects of the environment (reviewed in ref. 13).

Unlike its yeast relatives, however, *Neurospora* abhors repeated genes and other duplicated sequences in the genome. It has developed a mechanism, known as 'repeat-induced point mutation' (RIP), by which such repetitive sequences are destroyed through mutation¹⁴. It may have evolved this mechanism to keep out 'foreign' DNA from viruses and the like, with the result that very little redundancy appears in the text, as Galagan *et al.* show. Nearly every word has a distinct meaning — a feature that may help in decoding how the genome functions. When there is unknown redundancy, it presents a stumbling-block to those who attempt to understand the function of a gene by inactivating it.

So far the focus has largely been the study of one or a few words, but biologists can now begin to see larger syntactic units, such as the coincidence of the 'RIPed' text and its modification with methyl groups. In particular, Selker *et al.*⁵ have found that methylation occurs almost exclusively in the relics of transposons — mobile genetic elements — that have previously been inactivated by RIP. Methylation is generally associated with

gene silencing, and so RIP and methylation together seem to represent a belt-and-braces strategy to ensure that transposons remain inactive. However, Galagan and colleagues' intriguing finding that some regions of the text are methylated even though they are not RIPed raises the possibility that methylation also has another function in *Neurospora*.

Beadle and Tatum² said that, "From the standpoint of physiological genetics the development and functioning of an organism consist essentially of an integrated system of chemical reactions controlled in some manner by genes". With the *Neurospora* genome in hand, researchers can now move towards a specification of the chemical reactions that link genes and proteins in processes such as metabolism, biological clocks and development¹⁵. We may also be able to borrow from computer science to describe the reaction network in terms of a 'biological circuit' and discover more about how this fungus functions. New genomics approaches can measure the responses of the circuit (through RNA and protein profiling) and the links between its components (through protein-protein and protein-DNA

interaction mapping). The genetic text, together with such approaches, should bring a further revelatory consilience of biochemistry and genetics¹⁶.

Jonathan Arnold is in the Department of Genetics and Nelson Hilton is in the Department of English, University of Georgia, Athens, Georgia 30602-7223, USA.

e-mails: arnold@arches.uga.edu
nhilton@arches.uga.edu

1. Davis, R. H. *Neurospora: Contributions of a Model Organism* (Oxford Univ. Press, 2000).
2. Beadle, G. W. & Tatum, E. L. *Proc. Natl Acad. Sci. USA* **27**, 499–506 (1941).
3. Raju, N. B. *Eur. J. Cell Biol.* **23**, 208–223 (1980).
4. Galagan, J. E. *et al. Nature* **422**, 859–868 (2003).
5. Selker, E. U. *et al. Nature* **422**, 893–897 (2003).
6. Kelkar, H. S. *et al. Genetics* **157**, 979–990 (2001).
7. Kraemer, E., Wang, J., Guo, J., Hopkins, S. & Arnold, J. *Bioinformatics* **17**, 901–912 (2001).
8. Adams, M. D. *et al. Science* **287**, 2185–2195 (2000).
9. Venter, J. C. *et al. Science* **291**, 1304–1351 (2001).
10. International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
11. Goffeau, A. *et al. Science* **274**, 546–567 (1996).
12. Dunlap, J. C. *Cell* **96**, 271–290 (1999).
13. Gustin, M. C., Albertyn, J., Alexander, M. & Davenport, K. *Microbiol. Mol. Biol. Rev.* **62**, 1264–1300 (1998).
14. Selker, E. U. *Adv. Genet.* **46**, 439–450 (2002).
15. Ideker, T. *et al. Science* **292**, 929–934 (2001).
16. Battogtokh, D., Asch, D. K., Case, M. E., Arnold, J. & Schuttler, H.-B. *Proc. Natl Acad. Sci. USA* **99**, 16904–16909 (2002).

Quantum electronics

The electron is cool

Peter E. Toschek

Temperature is an awkward concept for a single particle, but the energy of the particle's 'quivering' is a useful substitute. Feedback control of this motion can be used to cool a single electron to very low temperature.

Individual electrons, trapped by electric and magnetic fields, can be used for precise, fundamental measurements. D'Urso, Odom and Gabrielse¹ now report in *Physical Review Letters* that they have succeeded in controlling the residual random motion of a trapped electron using feedback signals. This degree of control could be the basis of experiments to measure elementary constants of nature with even greater precision than before, and to accurately test the quantum theory that governs electrons and radiation — quantum electrodynamics.

The 'calming' of isolated quantum systems, by reducing their kinetic energy, is behind much of the progress in the physics of light and atoms in the past decades. Laser cooling has produced much narrower spectral lines of emitted radiation from free atoms² and bound ions^{3–5}, and hence more precise atomic data⁶. Evaporative cooling of atom clouds in magnetic or optical traps — similar to the way in which coffee cools in a cup — made possible the phase transition to a new state of matter, a Bose-Einstein condensate^{7,8}. Experimenters can force particles such as antiprotons to condense into tight bunches inside the storage ring of

an accelerator by detecting their random excursions and feeding that information back in the form of a changing electrode voltage to coax them back into line. This reduction of random motion is stochastic cooling⁹ — a decisive factor in the discovery of the W and Z bosons, the fundamental particles that mediate nature's weak force.

Feedback, a controlling operation determined by the result of a preceding measurement, may be imposed on any system. In earlier experiments, the residual motion of a single ion inside an electrodynamic trap (consisting of a static and an alternating electric field) has been reduced using a kind of feedback^{10,11}. Here, sequences of almost-resonant light pulses absorbed quanta of vibrational energy from the ion, damping its fluctuations. Now D'Urso *et al.*¹ demonstrate that it is possible to cool an individual quantum particle, an electron, by straightforward, continuous electronic feedback.

In their experiment, an electron is captured and trapped in an electromagnetic cage created from a homogeneous magnetic field and an inhomogeneous electric field that is zero at the centre of the apparatus (Fig. 1). This technique was pioneered some

25 years ago in the Nobel-prizewinning work of Hans Dehmelt¹². The magnetic field curls the path of the trapped electron into a spiral-like trajectory. A positively charged torus and two negatively charged caps, aligned on the trap axis, prevent the electron escaping. The upper cap also serves as a sensitive probe for the tiny voltage fluctuations caused by the hovering electron. Through a feedback loop, those voltage changes are amplified and fed back to the lower cap, creating corrective tugs on the electron's motion.

The random voltage that is induced by the thermally fluctuating electron motion along the axis of the trap has been measured in a frequency-selective way. The result is a spectrum of noise power that has a notch at the frequency value that matches the resonant frequency of this axial electronic oscillation — the electron can be represented by a serial circuit that acts, on resonance, as an electric short between the caps. The width of the notch represents how much the fluctuating electron motion is damped, characterizing the dissipation of vibrational electron energy into the electric circuit. Increasing the signal gain in the feedback loop causes this damping to decline, and with it shrinks the width of the spectral notch (Fig. 1).

The temperature of an ensemble of particles quantifies the fluctuations of this system. When this indicator is applied to a single electron, the particle's motional excursions are measured again and again, and the stochastic ensemble of observations can be easily transformed into recorded values of motional energy; these snapshots form, in the equilibrium state, a thermal (or Boltzmann) distribution, with higher energy values being exponentially less likely than lower ones. A tricky way of monitoring such a distribution — called thermometry — has been suggested by Dehmelt^{13,14}. There are, apart from the electron's axial motion, two more, independent modes of vibration: a fast azimuthal motion in which the electron orbits the trap axis, called 'cyclotron motion'; and a slow pulsation of the orbital radius, called 'magnetron motion'. These motions may be excited in differently quantized portions of vibrational energy, and a temperature measurement amounts to finding out how often and for how long the first excited state (and the second, third, and so on) of the quantized motion is randomly occupied by the electron — the mean numbers of occupation forming a Boltzmann distribution when the electron is in thermal equilibrium.

The job of measuring this probability of excitation in a non-invasive and precise way would seem to require a complex strategy — and the patience of indefatigable Sisyphus. The elegant (although not quite straightforward) solution is to make frequency measurements, which are famously precise and convenient, using two coupled oscillators known to be shifting each other's

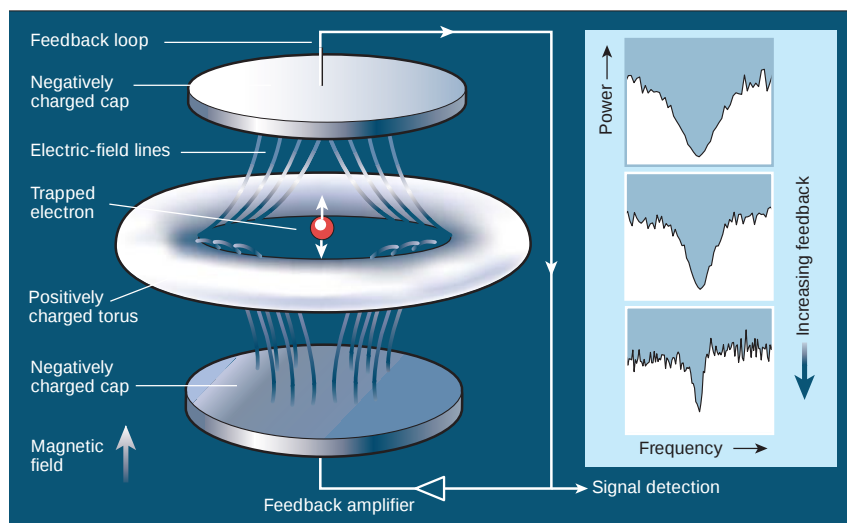


Figure 1 Cooling down. In the device designed by D'Urso *et al.*¹, a single electron is trapped between two caps and a torus by the combined effect of an electric and a magnetic field. The upper cap detects the random motion of the electron, and a signal sent through the feedback circuit triggers a voltage change on the lower cap to counteract it. The effective cooling of the electron is revealed in the power spectrum of noise in the circuit (inset; derived from ref. 1): a notch in the spectrum shrinks as the feedback increases; the frequency at which the notch appears is the resonant frequency of the electron between the caps.

frequency of oscillation in step with their own excitation, as two connected pendula do in the familiar classroom experiment. Here, it suffices to couple the cyclotron and the axial motion through the minute inhomogeneous magnetic field of a tiny nickel wire.

To measure the thermal distribution of the electron's energy of axial vibration requires a two-step 'pump-probe' operation. Because any axial energy value is correlated with a shifted cyclotron frequency, the probability of an axial excitation — its 'strength' — is mapped onto the particular strength of cyclotron excitation when the electron is irradiated with microwaves of a frequency correspondingly shifted off resonance. The quanta of cyclotron excitation are tiny compared with light quanta, but much larger than those of axial excitation: the electron absorbs at most one of them when randomly jumping to its first cyclotron-excited quantum state. The strength of excitation manifests itself in the fraction of successful quantum jumps to this state when a series of microwave 'pump' pulses is applied.

Then, to detect whether the electron has reached the first excited state after a pump pulse, the inverse correlation, namely of cyclotron energy and axial frequency, comes into play. A strong 'probe' pulse, at the axial frequency on resonance, does not affect the cyclotron-excited electron — except when the probe frequency is tuned off resonance by a tiny shift, just 12 Hz in D'Urso and colleagues' experiment¹. Thus, the mean chance of axial excitation is detected, via cyclotron excitation, by recording the response of the electron, seen by the electronic circuit, to the

probe pulse at the shifted axial frequency.

When the microwave pump radiation is varied in steps around the cyclotron resonance, the resulting spectrum of measurements shows how likely it is that the electron has acquired one quantum of cyclotron motion. This spectrum mimics a Boltzmann distribution, whose width directly displays the 'temperature' of the electron's axial vibration. The measurements¹ show that, when the strength of the feedback is increased, the width of the distribution shrinks down — the lowest temperature reached in their experiment was 850 mK.

Singling out and cooling electrons gives a means of measuring, with very great

accuracy, fundamental quantities in quantum electrodynamics, such as the ratio of the electron's spin to its magnetic moment. In a heroic effort by theorists, this number has been calculated to ten decimal places, and so far, no significant discrepancy has been found between it and the value measured in experimental tests. But it is quite possible that even higher precision may reveal a deviation — or even that the value for the electron's antiparticle twin, the positron, is different. This is crucial in the matter-antimatter relationship that is at the very heart of our understanding of matter. D'Urso and colleagues' feedback cooling may well contribute to dramatically reducing errors in these measurements, furthering our never-ending struggle for higher accuracy of observation and refined modelling of nature. ■

Peter E. Toschek is at the Institut für Laser-Physik, Universität Hamburg, Jungius-Straße 9, D-20355 Hamburg, Germany.

e-mail: toschek@physnet.uni-hamburg.de

1. D'Urso, B., Odom, B. & Gabrielse, G. *Phys. Rev. Lett.* **90**, 043001 (2003).
2. Hänsch, T. W. & Schawlow, A. L. *Opt. Commun.* **13**, 68–69 (1975).
3. Wineland, D. J., Drullinger, R. E. & Walls, F. L. *Phys. Rev. Lett.* **40**, 1639–1642 (1978).
4. Neuhauser, W., Hohenstatt, M., Toschek, P. E. & Dehmelt, H. G. *Phys. Rev. Lett.* **41**, 233–236 (1978).
5. Neuhauser, W., Hohenstatt, M., Toschek, P. E. & Dehmelt, H. G. *Appl. Phys.* **17**, 123–129 (1978).
6. Wineland, D. J. *et al.* in *Proc. 4th Symp. Freq. Standards Metrology* (ed. DeMarchi, A.) 71–77 (Springer, Berlin, 1989).
7. Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. *Science* **269**, 198–201 (1995).
8. Davis, K. B. *et al.* *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
9. van der Meer, S. *Rev. Mod. Phys.* **57**, 689–697 (1985).
10. Eschner, J., Appasamy, B. & Toschek, P. E. *Phys. Rev. Lett.* **74**, 2435–2438 (1995).
11. Appasamy, B., Stalgies, Y. & Toschek, P. E. *Phys. Rev. Lett.* **80**, 2805–2808 (1998).
12. van Dyck, R. S. Jr, Schwinberg, P. B. & Dehmelt, H. G. *Phys. Rev. Lett.* **38**, 310–313 (1977).
13. van Dyck, R. S. Jr, Ekstrom, P. & Dehmelt, H. *Nature* **262**, 776 (1976).
14. Dehmelt, H. *Am. J. Phys.* **58**, 17–27 (1990).

Stem cells

Fusion brings down barriers

Alexander Medvinsky and Austin Smith

It remains uncertain how tissue-specific stem cells could generate the mature cell types of another tissue. In one instance, where bone-marrow-derived stem cells repair damaged liver in mice, cell fusion is the answer.

The respective merits of embryonic versus adult stem cells for treating diseases have been widely discussed, in both the popular press and the scientific literature. In particular, reports that the developmental potential of adult stem cells might be greater than previously supposed have aroused strong interest and controversy. The papers by Wang and colleagues¹ and Vassilopoulos and co-workers² on pages 897 and 901 of this issue mark a further turn in

the debate. These authors confirm that stem cells derived from adult bone marrow can repair damaged liver tissue in mice — but not by converting directly into liver cells, as might be expected if the stem cells could change their developmental 'destiny'. Instead, cell fusion with the host liver is responsible for bringing down the barriers between bone marrow and liver.

Many tissues and organs in adult mammals contain reserves of stem cells to ensure

their long-term anatomical and functional maintenance. Stem cells are very few in number, but have both a high proliferation potential — allowing their lifelong renewal — and the capacity to generate fully mature, tissue-specific cell types. These strategic cell reserves are recruited in response to physiological demands or tissue damage.

Adult organs have generally been considered to be closed cell communities, reflecting their distinct developmental origins. Thus it was assumed, for instance, that stem cells from the brain are restricted to producing brain-specific mature cell types. Recently, however, the dogma of 'tissue fidelity' has been challenged. There have been reports that, following transplantation, cells derived from bone marrow can turn into liver, muscle, nerve and other specialized cell types. Conversely, it has been claimed that stem cells from the brain can produce blood or muscle. Such reports have given rise to the concept that stem cells from adult tissues are not constrained by their ancestry, and have the 'plasticity' to alter their destiny³.

But some investigators have been sceptical about these findings⁴. A key issue is the true identity of cells produced through 'tissue conversion'. To show that a certain cell type originated from a transplanted stem cell, the presence of both a marker gene from the donor and a tissue-specific protein must be demonstrated. But to show this unambiguously is technically demanding. And it is

even more difficult to show that a particular cell, even if it appears to be specialized, actually functions in a tissue-specific way. In addition, 'converted' cells often occur singly, a fact that is difficult to reconcile with a stem-cell origin, which is expected to generate clusters of progeny.

These concerns appeared to be addressed in a study showing that mice with a fatal metabolic liver disease — a vital enzyme, fumarylacetoacetate hydrolase (Fah), is missing in these animals — can be 'rescued' by the transplantation of purified blood stem cells⁵. Histological examination revealed massive areas of healthy, donor-derived liver tissue. This observation was reproducible and the livers were demonstrably functional, as the animals survived. So blood can produce liver — but how? Do the blood stem cells directly generate liver cells (Fig. 1a)?

It seems not. A well-known way by which cells can change identity is through cell fusion⁶. In hybrid cells produced by fusion, molecules from one fusion partner reprogramme gene expression in the genome of the other partner. Fusion can occur spontaneously *in vitro* — this was shown for the first time in the 1960s, and again more recently in reports^{7,8} that also indicated that hybrid cells can function *in vivo*. Wang *et al.*¹ and Vassilopoulos *et al.*² now show unequivocally that the cells that build healthy liver tissue in Fah-deficient mice contain both donor and host genetic markers. Cells with such a

genetic constitution can only have arisen by *in vivo* fusion. So, the donor cells effect functional rescue by delivering their genome, which includes the Fah gene, into pre-existing liver cells (Fig. 1b). In the resulting hybrids, liver-cell molecules dominate over blood-cell factors, so that liver gene expression is activated and blood gene expression is silenced.

The ancients recounted the myth of Prometheus, whose damaged liver was continuously regenerated. In this tale — and in liver regeneration in general — it is not stem cells that are required, but rather the proliferation of mature liver cells. In the new work^{1,2}, the contribution of the hybrid cells seems to depend on this regenerative context, and on the space created by degeneration of the diseased host liver. (This may explain why, in healthy liver, transplanted bone-marrow cells yield only single liver cells⁹.) The liver may also be a particularly favourable setting for hybrid cells, because normal liver cells often carry several copies of each chromosome; hybrids, too, generally have an additional set of chromosomes (Fig. 2).

One important remaining question is exactly which cells fuse with the host liver; there is no evidence that it is the stem cells themselves. Instead, it seems more likely that differentiated progeny of the stem cells, such as blood cells known as macrophages, are responsible, because a contribution to the liver is seen only after the donor stem cells have populated the animals' blood system^{1,2}. It will be essential to identify the precise fusion partners in order to optimize any clinical applications. Another question is whether the Fah gene is already active in certain donor-derived blood cells, or whether it is activated after fusion. Furthermore, the extent to which gene expression throughout the genome is reprogrammed in hybrids — and whether this is prone to error, as is the case with cloning — will require meticulous investigation.

Another remarkable observation reported by Wang *et al.*¹ is that some cells in the rescued liver contain only two copies (a pair) of each chromosome. How could fusion occur without an increase in chromosome number? The most probable explanation is that hybrid cells undergo a 'reduction division', in which an entire set of paired chromosomes is lost. Such reduction divisions will conceal fusion history (Fig. 2), introducing a further confounding factor to the investigation of tissue conversion.

The issue of stem-cell plasticity in its pure form — that is, whether a stem cell from one tissue can transform into a different tissue type — will remain open for some time. Last year's discovery of multipotent adult progenitor cells (MAPCs)¹⁰ added a further dimension. These cells can produce various mature cell types in isolation, and therefore without fusion. But how do MAPCs arise in

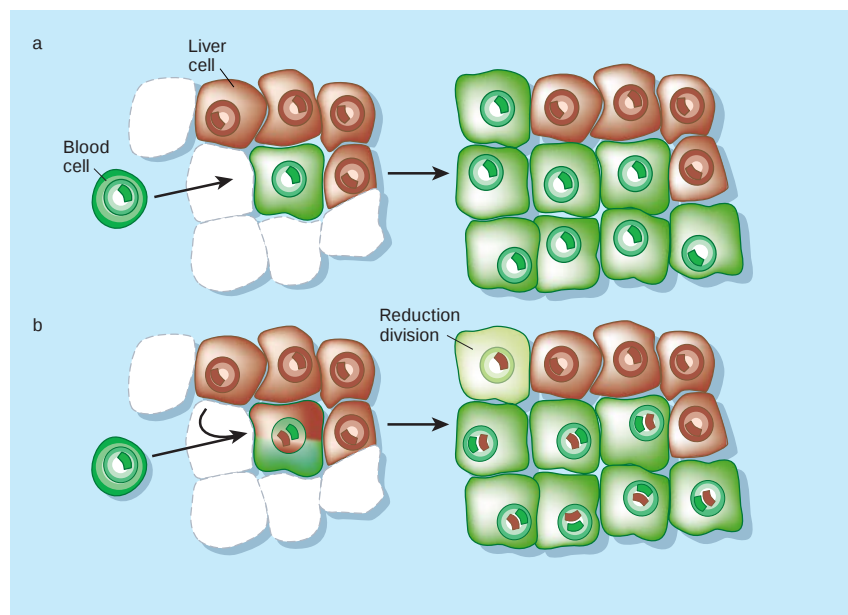


Figure 1 Possible mechanisms by which transplanted blood cells could transform into liver cells. **a**, Direct conversion of donor blood cells into liver cells, induced by the liver environment. A converted cell and its daughters contain only donor chromosomes. (Marker chromosomes are depicted, by which donor, green, and host, brown, can be discriminated.) **b**, Fusion of donor blood cells with resident liver cells. A hybrid cell and its progeny generally contain both donor and host chromosomes. The indicated hybrid cell underwent reduction division, reducing the number of chromosome pairs to the norm (see Fig. 2). The studies by Wang *et al.*¹ and Vassilopoulos *et al.*² demonstrate scenario **b**.

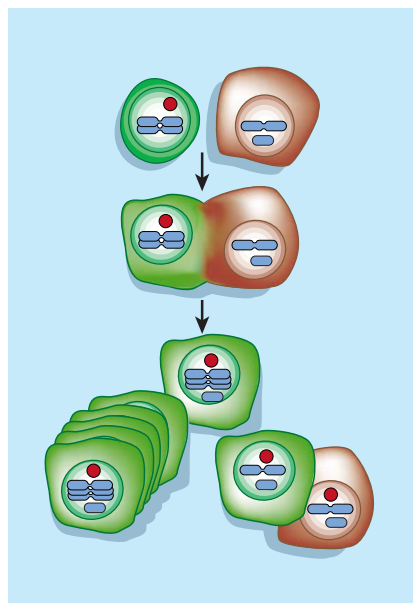


Figure 2 Formation and reduction of hybrids. Cell fusion first produces a cell containing two nuclei; the exchange of specific molecules reprogrammes gene expression. The nuclei then fuse to generate a mononuclear hybrid cell that is polyloid (it has more than the usual chromosomal content). The hybrid divides to give identical polyloid daughters (lower left). Wang *et al.*¹ present evidence that hybrids can also generate daughters of normal ploidy, presumably through a reduction division (lower right). The expulsion of chromosomes can mask the fusion history of these cells. Blue shading indicates chromosomes. Red represents a specific gene product from the donor (in this case, the *Fah* enzyme).

the first place? They seem to develop over time in culture, possibly because they become liberated from tissue-restricted gene expression. Could it be that fusion between distinct tissue-specific cell types in the starting population creates hybrids that reprogramme to become MAPCs?

Our new understanding of how blood cells produce liver^{1,2} should have an impact on research into the use of bone-marrow transplantation for treating certain genetic disorders: gene transfer through *in vivo* fusion now seems a distinct possibility. But the major applications of regenerative medicine seem likely to be delivered through embryonic and tissue-specific adult stem cells.

Alexander Medvinsky and Austin Smith are at the Institute for Stem Cell Research, University of Edinburgh, King's Buildings, Edinburgh EH9 3JQ, UK.
e-mails: a.medvinsky@ed.ac.uk
austin.smith@ed.ac.uk

1. Wang, X. *et al.* *Nature* **422**, 897–901 (2003).
2. Vassilopoulos, G., Wang, P.-R. & Russell, D. W. *Nature* **422**, 901–904 (2003).
3. Blau, H. M., Brazelton, T. R. & Weimann, J. M. *Cell* **105**, 829–841 (2001).

4. Anderson, D. J., Gage, F. H. & Weissman, I. L. *Nature Med.* **7**, 393–395 (2001).
5. Lagasse, E. *et al.* *Nature Med.* **6**, 1229–1234 (2000).
6. Ephrussi, B. *Hybridization of Somatic Cells* (Princeton Univ. Press, 1972).

7. Terada, N. *et al.* *Nature* **416**, 542–545 (2002).
8. Ying, Q. L., Nichols, J., Evans, E. P. & Smith, A. G. *Nature* **416**, 545–548 (2002).
9. Wagers, A. J., Sherwood, R. I., Christensen, J. L. & Weissman, I. L. *Science* **297**, 2256–2259 (2002).
10. Jiang, Y. *et al.* *Nature* **418**, 41–49 (2002).

Astronomy

Telling the tale of the first stars

Timothy C. Beers

HE0107–5240 is a star in more than one sense of the word. Chemically, it is the most primitive object yet discovered, and it is at the centre of debate about the origins of the first elements in the Universe.

Late last year, the discovery of the most iron-deficient star yet identified, HE0107–5240, was announced. This star has a measured abundance of iron less than 1/200,000 that of the Sun¹. Its significance is that it seems to be a relic from the early Universe, and astronomers are now busy considering how to interpret it.

In this issue, three groups — Bonifacio *et al.*², Umeda and Nomoto³, and Schneider *et al.*⁴ — present various interpretations of HE0107–5240. Each of these contributions centres on whether this star exhibits properties that might reveal the likely range of mass that should be associated with the so-called population III stars — objects that are presumed to have formed shortly after the Big Bang, and which are thought to have produced the first elements heavier than H, He and Li, as well as the ‘first light’ in the Universe. Population II stars are objects that formed after population III stars, and which incorporated the metals created by this previous generation. Our Sun, and other (younger) metal-rich stars in the Galactic disk, are referred to as population I objects.

To the astronomer, metals include all elements heavier than He, and they are thought to be produced only by nuclear reactions that take place during the lifetimes, or at the deaths, of stars. Stars such as our Sun have inherited the net production of metals by all of the previous generations that lived (and died) before it. Stars with the lowest observed abundances of heavy elements, such as HE0107–5240, must therefore have been born before other stars, because the gas clouds from which they formed had only the slightest traces of these heavy elements. So, regardless of the outcome of debate about the nature of the very first stars, HE0107–5240 remains chemically the most primitive object yet discovered, and is a crucial ‘laboratory’ for tests of the origins of the first elements in the Universe.

Despite the apparent simplicity of the composition of HE0107–5240, which to date has yielded the detection of nine elements (H, C, N, Na, Mg, Ca, Ti, Fe and Ni), the new papers^{2–4} present a dizzying

array of possible explanations for their origin. By comparison, other extremely low-metallicity stars, with Fe abundances near 1/1,000 the solar level, such as CS22892–052 (ref. 5) and CS31082–001 (refs 6, 7), show evidence of roughly 40–60 individual elements, most arising from the so-called rapid neutron-capture process, which accounts for the production of roughly half of all elements heavier than Fe. Beyond its Fe deficiency, the singular feature of HE0107–5240 is that its measured abundance of C, relative to Fe, is about 10,000 times the observed ratio of these elements in the Sun, the largest such ‘over-abundance’ ratio ever seen. The N abundance ratio is also greatly enhanced, though only by a factor of 200. The other detected elements exhibit ratios similar to those in previously identified metal-deficient stars.

Explanations put forward for the composition of HE0107–5240 fall into three main categories. First, that it is indeed a low-mass, population III star that formed out of gas of zero metallicity, and has had its present surface abundances altered. Possible mechanisms for changing the observed atmospheric abundances of HE0107–5240 include internal mixing of elements produced by nuclear burning in the star itself, and the acquisition of metals produced by later generations of stars during its passage through an already enriched interstellar medium⁸, or — as now seems more likely — even in the cloud of its birth, from the contribution of material from later supernovae⁹.

The second possible explanation is that it is a low-mass population II star, of an extreme form, and that the observed elemental abundances directly reflect the yields of species from supernova explosions of massive (more than 200 times the solar mass) population III stars that were incorporated into the gas from which HE0107–5240 formed.

Third, as above, except that the observed abundances reflect the yields of one or more supernova explosions of population III stars of ‘normal’ mass (20–25 solar masses) that were present shortly before its birth.

An important point here is that each of

Ecology

A liverwort cheat

Ghostwort — *Cryptothallus mirabilis* — is aptly named: it is a liverwort that lives beneath the surface layers in woodland, and rarely comes to human attention. Martin I. Bidartondo and colleagues tell how they have delved into the tangled details of its relationships with other organisms, and brought to light its cheating way of life (*Proc. R. Soc. Lond. B* **270**, 835–842; 2003).

Cryptothallus cannot carry out photosynthesis, and so must have a different energy source. It has long been known that it is associated with certain fungi. In a chain of experiment

and inference, in part involving growth of the various players in microcosms, Bidartondo *et al.* now find that the fungi concerned belong to the *Tulasnella* group which, in turn, form mycorrhizae — close and mutually beneficial connections (pictured) — with the roots of trees such as birch and pine.

So far, so cosy. But from work with carbon isotopes it turns out that *Cryptothallus* is a cheat: it gets its carbon supply not from the soil, as once thought, but from the tree via *Tulasnella*. Instances of such complex relationships are



known from other plants and other fungi, but this example greatly widens the field of organisms that can be involved.

Tim Lincoln

these options is consistent with the possibility that HE0107–5240 has ‘locked up’ the elemental clues required for understanding the nature of population III stars. Astronomers have long debated the form of the so-called ‘initial mass function’ of these objects, although recent theory has favoured the notion that it was dominated by stars of between several hundred and one thousand times the mass of the Sun. Such massive stars have extremely short lifetimes, so the elements they created provide one of the few lingering pieces of evidence that can be used to infer their properties.

Detailed observations of HE0107–5240, now under way, should help to discriminate between the three options. On page 834, for instance, Bonifacio *et al.*² propose that measurement of the element O may hold one key. They argue that this element should be detectable in the near-ultraviolet region of the spectrum of HE0107–5240. If it turns out to show a ratio, with respect to Fe, of more than about 1,000, this may be associated with production from high-mass, zero-metallicity stars. A value of less than 1,000 would suggest an origin from zero- (or low-) metallicity, lower-mass stars.

Bonifacio *et al.* also suggest that the presently observed metals in HE0107–5240 might have arisen from the yields of two supernova explosions of zero-metallicity progenitors with quite different masses — the high-mass progenitor producing the light elements now observed in the star (but little or none of the Fe), and the low-mass progenitor contributing the elements Mg and heavier (including Fe). Similar ideas, calling for a combination of supernovae to explain the observed elemental abundances of metal-poor stars, have already been proposed¹⁰.

Umeda and Nomoto (page 871)³ argue that the observed abundance pattern in HE0107–5240, including the large abundances of C and N, can be explained by the explosions of 20–130-solar-mass progenitors that underwent substantial ‘mixing and fallback’. According to this view, most of the heavy elements (including Fe) that were created in the central regions of explosions were never ejected, but fell into the maw of a black hole that formed at the death of the progenitor. The lighter elements, such as C and N, which formed by pre-explosion nucleosynthesis in the progenitor, were ejected into the interstellar medium. Umeda and Nomoto also note that similar conditions might arise from aspherical supernova explosions, where jets of explosive nucleosynthesis might have been produced by a rapidly rotating progenitor star. These ideas are supported by the existence of other C-, N- (and O-) rich, extremely metal-poor stars, which have abundance patterns that are in some ways like that of HE0107–5240 and could be explained by a similar mechanism. Furthermore, the sort of low-luminosity supernova explosions envisaged by Umeda and Nomoto have already been observed — examples include SN1997D and SN1999br.

Both Umeda and Nomoto³ and Schneider *et al.*⁴ (page 869) point out that the great abundance of C, N (and presumably O) in HE0107–5240 may indicate that the gas cloud from which it formed was already quite metal-rich (even though the Fe abundance is low, the total abundance of metals is dominated by the lighter elements). That would allow efficient formation of low-mass stars from the ‘cooling channels’ supplied by the lighter elements. But these authors also

note that, even if the abundances of C, N and O in the cloud from which HE0107–5240 formed were quite low, and were boosted only later from internal mixing processes, gas with metal abundance as low as that inferred from the Fe alone could still fractionate to form low-mass stars due to cooling from dust grains — if indeed such grains could have formed at sufficiently early times.

Schneider *et al.*⁴ also argue that if the first population III stars had masses of 200–220 times the solar mass, their explosions might account for the abundances of the heavier elements in HE0107–5240, though not the lighter ones. Schneider *et al.* agree with Umeda and Nomoto that another possibility is the explosion of progenitors with masses 20–25 times that of the Sun, and point out that improved upper limits on the abundance of Zn, or the presence (or lack) of elements created in the rapid neutron-capture process in HE0107–5240, may be able to discriminate between the appropriate mass range of the progenitor object. From the data already in hand, however, Christlieb *et al.*¹ have argued that if the neutron-capture elements were indeed greatly enhanced, Ba and Sr might be expected to be detected in this star. But they are not.

Although we are left with a frustrating variety of possibilities, study of HE0107–5240 should allow us to address many questions about the formation and evolution of the first generations of stars in the Universe. One line of attack, not explicitly mentioned in these three papers, is analysis of the metallicity distribution function of stars with the lowest abundances of heavy elements⁹. If they apply in general, several of the explanations offered for the formation of a star with the observed properties of HE0107–5240 will produce stars with characteristic metallicities — which, given a large enough sample, might be detected as deviations from a continuous distribution of stellar metallicities. Hence, numerous additional stars with extremely low Fe abundances will need to be discovered to fully ‘tell the tale’ of early star formation and the creation of the first metals in the Universe.

Timothy C. Beers is in the Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA.
e-mail: beers@pa.msu.edu

- Christlieb, N. *et al.* *Nature* **419**, 904–906 (2002).
- Bonifacio, P., Limongi, M. & Chieffi, A. *Nature* **422**, 834 (2003).
- Umeda, H. & Nomoto, K. *Nature* **422**, 871–873 (2003).
- Schneider, R., Ferrara, A., Salvaterra, R., Omukai, K. & Bromm, V. *Nature* **422**, 869–871 (2003).
- Snedden, C. *et al.* *Astrophys. J.* (in the press); astro-ph/0303542 (2003).
- Cayrel, R. *et al.* *Nature* **409**, 691–692 (2001).
- Hill, V. *et al.* *Astron. Astrophys.* **387**, 560–579 (2002).
- Yoshii, Y. *Astron. Astrophys.* **97**, 280–290 (1981).
- Shigeo, T., Tsujimoto, T. & Yoshii, Y. *Astrophys. J.* **586**, L57–L60 (2003).
- Wasserburg, G. J. & Qian, Y.-Z. *Astrophys. J.* **529**, L21–L24 (2000).

Cancer

Escape from inhibition

Sara A. Courtneidge

Mutations in the Abl protein cause some leukaemias. The determination of its structure and the identification of Abl variants resistant to an anti-leukaemia drug reveal regions of the protein needed for its regulation.

The enigmatic protein Abl has just given up another of its secrets. This enzyme is important because in mutated, activated form it causes some types of human leukaemia — so researchers have been keen to learn how it is regulated. That goal became still more precious when it was realized that the resistance of leukaemia to Gleevec (a drug that inhibits Abl) often results from further mutations in Abl itself. But until now the protein has stubbornly resisted attempts to understand it.

Writing in *Cell*, however, Hantschel and colleagues¹ and Nagar and co-workers² describe how they used a combination of mutational and crystallographic analyses to solve the puzzle. In a surprising twist, the regulation of Abl turns out to closely resemble that of a related enzyme, Src, despite key structural differences. Meanwhile, a complementary paper by Azam and colleagues³ in the same issue describes the use of an *in vitro* assay to screen for Gleevec-resistant variants of Abl. This work reveals that Abl can escape Gleevec inhibition by mutating several amino acids now shown to be key to its regulation. Together these findings have implications for the design of the next generation of Abl inhibitors, and provide an unbiased assay with which drug resistance in Abl can be probed in the future.

Abl is a tyrosine kinase — an enzyme that transfers phosphate groups to tyrosine amino acids in proteins — and is expressed in all tissues of the body. One end of the protein (the amino-terminal half) resembles the related kinases from the Src family⁴, with a region for binding proline-rich peptides (the SH3 domain), a domain that binds

proteins containing phosphorylated tyrosine residues (the SH2 domain), and a catalytic portion (Fig. 1). But whereas Src-family kinases have, following the catalytic domain, a short 'tail' (which is crucial for their regulation), Abl has a long extension with domains that dictate its subcellular localization, and regions for protein and DNA binding⁵. Alternative processing of the Abl-encoding messenger RNA produces two proteins, called type 1a and type 1b, which differ in their extreme amino termini. Both forms of Abl are tightly regulated in cells, generally have very low kinase activity, and are distributed between the cell cytoplasm and nucleus.

So how is Abl so tightly controlled? One reason why researchers are interested in this question is because deregulation of Abl's kinase activity causes certain forms of blood cancer. More than 95% of chronic myelogenous leukaemias (CMLs), and some 10% of acute lymphocytic leukaemias, result from a chromosomal translocation that fuses part of the *bcr* gene from chromosome 22 with part of the *abl* gene on chromosome 9. The result is a fusion protein, Bcr-Abl⁶, that is missing the extreme amino terminus of Abl (Fig. 1) and has high kinase activity, although it spends some time in the inactive state. The drug Gleevec (also known as STI-571 or imatinib), recently approved to treat CML, inhibits the catalytic activity of Bcr-Abl by binding to the inactive form of the kinase, thus stabilizing it^{7,8}. It is a very successful treatment for patients in the early, chronic stage of the disease. But most patients in the later stage of the disease, although responding well at first, eventually

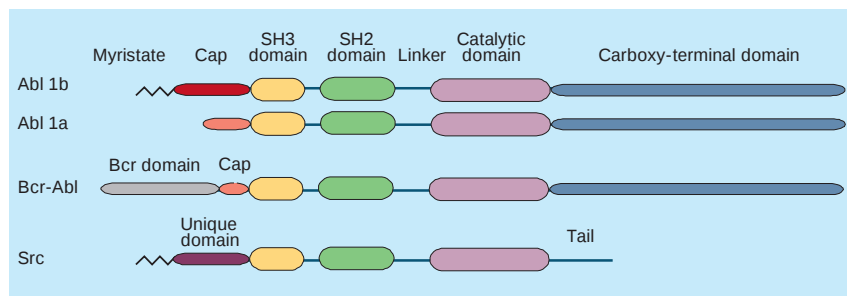


Figure 1 The domain structure of the enzymes Abl and Src. The domains of Abl 1a and 1b (the two alternatively processed forms of Abl), Bcr-Abl and Src are shown. Note that the Bcr domain is not depicted in its entirety. The Abl carboxy-terminal domain contains sites that can be phosphorylated, sequences for import into and export from the nucleus, motifs with the sequence PxxP (where P is proline and x is any amino acid), a DNA-binding domain and a region that binds the cytoskeletal protein actin. These have been omitted for simplicity.



100 YEARS AGO

The trouble of compiling pedigrees and their unmanageable size led me to devise a method of recording relationships in a form suitable to my own particular wants. As it promises to answer exceedingly well, and to be of more extending utility, I venture to publish it. The system of relationships between those who live or have lived in a long-established community is wide in extent, of indefinite depth, and interlaced in all directions. The problem is how to arrange its records so that when any individual is selected as a point of departure, it shall be easy to trace his relationship in every direction, whether ascending, descending, or collateral, so far as materials exist. The representation of such a system is wholly beyond the powers of a chart, but its object can be attained by breaking it up into what will be called "Family Groups," each of which slightly overlaps those with which it is immediately connected. A family group, in the sense used here, consists of (1) a parental couple, (2) all their sons and daughters, (3) the wives and husbands of them. Their names are supposed to be written on one page of a register, and the group, as a whole, to be defined by the No. of that page. The group is also... indexed under the joined surnames of the parental couple. Francis Galton
From *Nature* 23 April 1903.

50 YEARS AGO

New terms of grant were offered to nine research associations for the next five-year period... The new block grant for the British Hat and Allied Feltmakers' Research Association for 1952-54 will be £7,000 a year, conditional on £10,000 from the industry, with up to £5,000 on the £100 for £100 basis. The Association reports advances in the knowledge of mercuric carotting of rabbit fur, which are leading to a better understanding of the process, and an investigation has been started on the effect of different types of dyeing machine on the quality of felt produced. For the two years 1952-53, the British Food Manufacturing Industries Research Association will receive a block grant of £9,000 a year, conditional on £20,000 from industry, with up to a further £6,000 on the customary basis... the Association is investigating the graining and rheological behaviour of chocolate, which are important in controlling the physical character of the chocolate and in preventing 'bloom'.
From *Nature* 25 April 1953.

relapse⁹. Some relapses are due to mutations in the catalytic domain of Bcr–Abl. So, understanding the effect of these mutations requires that we understand how Abl is regulated under normal circumstances.

The regulation of Src, an enzyme closely related to Abl, is well understood (Fig. 2a). In the inactive form, a phosphorylated tyrosine in the tail of the Src protein binds to the protein's own SH2 domain. This triggers the association of the SH3 domain with sequences in the 'linker' that lies between the SH2 domain and the catalytic domain, which in turn forces the catalytic domain to adopt a conformation incompatible with catalysis¹⁰. Another domain of Src, the 'unique domain', has no role in its regulation, but is required to attach the protein to membranes (with linkage of the small membrane lipid myristate to the first amino acid of the unique domain being particularly important).

Given the similarities between Abl and Src, why was the regulation of Abl so difficult to understand? The answer lies partly in the fact that Abl lacks a tyrosine-phosphorylated tail, and indeed is not detectably tyrosine-phosphorylated in cells. Furthermore, mutation of Abl's SH2 domain (so that it can no longer bind phosphorylated tyrosines) does not activate the enzyme. This, together with the observation that Abl's SH3 domain is required for regulation, led to models that invoked not 'self-regulation' (as for Src), but rather regulation by other proteins binding to the SH3 domain⁵.

Prompted by two key observations, Hantschel *et al.*¹ and Nagar *et al.*² have now proposed a new model, in which regulation is in fact an intrinsic property of the amino-terminal half of Abl. The first observation was that sequences at the extreme amino-terminus of Abl 1b, called the cap, are critical^{1,11}. In particular, mutation of the first amino acid (which prevents attachment of myristate) activates the enzyme both *in vitro* and *in vivo*, without affecting its subcellular localization. Second, incubation of Abl with high concentrations of tyrosine-phosphorylated peptides activates catalysis, suggesting a role for the SH2 domain¹.

Analysis of several crystal structures of Abl 1b, with and without the cap and/or catalytic inhibitors, reveals how the cap and SH2 domains contribute to the self-inhibition of Abl² (Fig. 2b). Remarkably, the trigger that generates the inhibited state is the binding of the myristate-attached cap to a deep pit at the base of the catalytic domain, causing a conformational change in this domain. This in turn forces the SH2 domain to dock onto the catalytic domain in a way that prevents SH2 interacting with proteins containing phosphorylated tyrosines. As in Src, the SH3 domain, rigidly linked to the SH2 domain, makes contact with linker sequences. Despite the overall similarity between the inhibited forms of Abl and Src, there are some key

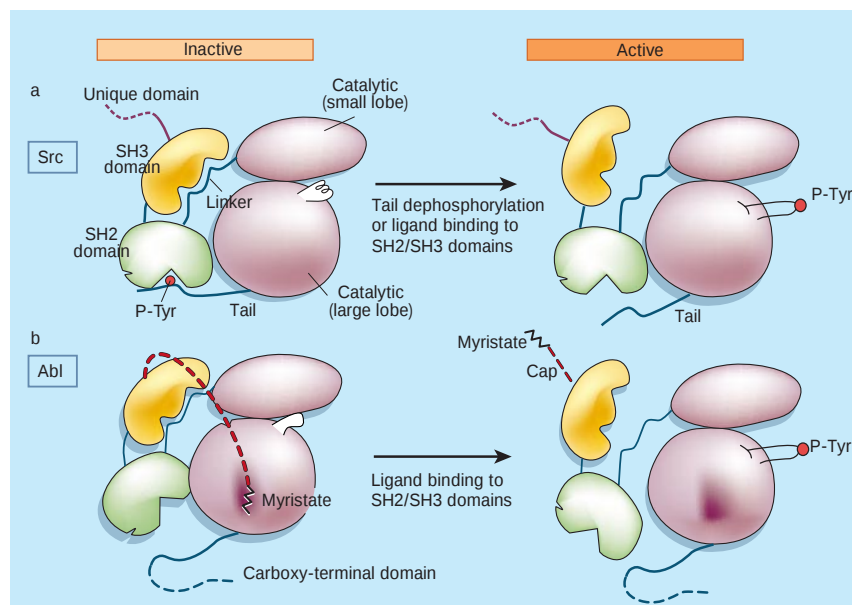


Figure 2 The regulation of Src and Abl. **a**, The inactive form of Src is initiated by an intramolecular interaction between the SH2 domain and the tyrosine-phosphorylated tail (P-Tyr), and also involves the SH3 domain and linker. **b**, Abl is locked into an inactive state by binding of the myristate-attached cap to a pit in the large lobe of the catalytic domain. The SH2 domain binds the catalytic domain in a way that occludes the SH2 domain's phosphotyrosine-binding site. The substrate-binding face of the SH3 domain interacts with the linker. These locking mechanisms cause the catalytic domains of Abl and Src to adopt (different) inactive conformations. Activation of Abl occurs when high-affinity ligands for the SH3 and/or SH2 domains break the intramolecular contacts. Activation of Src can also occur in this way, and by dephosphorylation of the tail. Note that except for the myristate group, the cap sequences were not visible in the latest crystal structure of Abl² (hence the dotted lines). Neither the carboxy-terminal domain of Abl nor the unique domain of Src was present in constructs used for structural analysis.

differences, particularly in the conformation of the catalytic domain, that help to explain why Gleevec inhibits Abl but not Src.

Which brings us back to the issue of resistance to Gleevec. In searching for mutations in Bcr–Abl that confer such resistance, most studies have concentrated on the catalytic domain. Azam *et al.*³ have conducted a more comprehensive survey, by testing randomly mutated Bcr–Abl proteins *in vitro* to identify Gleevec-resistant variants. Among the variants generated, the authors found most of the known catalytic-domain mutations that occur in patients, but also several others that affect the non-catalytic domains. Many of these mutations are in sites that the crystal structure now shows to be required in maintaining the inactive form of Abl—that is, in the cap, the SH3 and SH2 domains, and the linker. In these variants, the inhibited form of Abl is probably destabilized, so that the active enzyme, which cannot bind Gleevec, predominates.

In sum, we now understand that the cap, the SH3 domain and the SH2 domain are required both for regulation of Abl under normal circumstances, and for inhibition of Bcr–Abl by Gleevec. It is likely that, as Gleevec use expands, mutations outside the catalytic domain will also be detected in some drug-resistant patients. New Abl inhibitors that are being developed to treat Gleevec-resistant

patients should in the future be screened to see if they work on the most common of these mutant forms. Furthermore, Bcr–Abl is not attached to myristate; future drug-discovery efforts might therefore fruitfully be geared to identifying compounds able to bind the myristate-attachment site and restore full regulation to the enzyme.

But Abl has not given up all of its secrets yet. Although the structural analysis² revealed the myristate group in the pit of the catalytic domain, it was not possible to see the relationship between the rest of the cap and Abl. Also, the alternatively processed form Abl 1a is fully regulated even though its cap lacks myristate; and Bcr–Abl only has a partial cap. How a non-myristylated cap contributes to Abl regulation will surely be the subject of intense study.

Sara A. Courtneidge is at the Van Andel Research Institute, 333 Bostwick Avenue NE, Grand Rapids, Michigan 49503, USA.

e-mail: sara.courtneidge@vai.org

- Hantschel, O. *et al.* *Cell* **112**, 845–857 (2003).
- Nagar, B. *et al.* *Cell* **112**, 859–871 (2003).
- Azam, M., Latek, R. R. & Daley, G. Q. *Cell* **112**, 831–843 (2003).
- Superti-Furga, G. & Courtneidge, S. *BioEssays* **17**, 321–330 (1995).
- Smith, J. M. & Mayer, B. J. *Front. Biosci.* **7**, d31–d42 (2002).
- Mauro, M. J. & Drucker, B. J. *Curr. Opin. Oncol.* **13**, 3–7 (2001).
- Nagar, B. *et al.* *Cancer Res.* **62**, 4236–4243 (2002).
- Schindler, T. *et al.* *Science* **289**, 1938–1942 (2000).
- McCormick, F. *Nature* **412**, 281–282 (2001).
- Mayer, B. J. *Curr. Biol.* **7**, R295–R298 (1997).
- Pluk, H., Dorey, K. & Superti-Furga, G. *Cell* **108**, 247–259 (2002).

Cancer

Cell invasions and oestrogen receptors

Cell **113**, 207–219 (2003)

The cells of many breast cancers contain receptors for oestrogen and are strongly dependent on oestrogen for growth. But low levels of oestrogen receptors may be dangerous too. Naoyuki Fujita and colleagues suggest that a shortage of these receptor proteins could increase the likelihood that breast cancer will spread.

The ability of cancer cells to spread depends in part on their synthesis of cell-adhesion molecules. Such molecules help to anchor healthy cells in place; but in some breast cancers with poor prognoses, the levels of adhesion molecules are reduced. Fujita *et al.* now find that one such molecule, E-cadherin, is indirectly regulated by a protein called MTA3, and that synthesis of this protein is in turn linked to oestrogen. MTA3 is found only in cell lines that contain oestrogen receptors, and if receptor levels are artificially increased, MTA3 levels rise.

More worryingly, if oestrogen-receptor levels are lowered, MTA3 and E-cadherin levels drop — a condition associated with the invasive growth of breast cancers. This has implications for women with breast cancer who are taking selective oestrogen-receptor modulators. Any compound that interferes with the receptor's ability to activate MTA3 would be predicted to lower E-cadherin levels, and might predispose women to invasive tumours.

Helen R. Pilcher

Biomechanics

With a swing

J. Exp. Biol. **206**, 1631–1642 (2003)

The gibbon's swing has a built-in safety margin, according to James R. Usherwood and John E. A. Bertram. By putting more energy into each swoop than is needed to reach the next branch, they reduce the chances of a potentially fatal fall.

The most efficient way for a gibbon to travel would be like a smooth pendulum, swinging in arcs from branch to branch with little loss of energy in between. But Usherwood and Bertram found that this is not the case. In studies of gibbon movement around an artificial environment, they show that the animal has a built-in overshoot, in effect colliding with a branch and then accelerating into the next swing.

Gibbons (*Hylobates lar*) move their bodies and limbs to minimize the energy loss, it seems, and the apes' long arms also ease the jolt of deceleration. Gibbons have two 'gaits': one where they always keep one

hand on the branch, and another used for long distances between branches or high speeds, where they become airborne between handholds. Both involve overshooting. But the technique evidently isn't foolproof, because one-third of all gibbon skeletons show fractures. John Whitfield

Astrobiology

A salty life on Europa?

J. Geophys. Res. doi:10.1029/2002JE001966 (2003)

The ice-covered ocean of Jupiter's moon Europa is perhaps the best candidate habitat for extraterrestrial life in the Solar System. Mikhail Y. Zolotov and Everett L. Shock say that, at least early on in the moon's history, the ocean probably contained enough of the right ingredients to provide an energy source for living organisms.

Satellite observations show that Europa's hidden ocean is salty, thanks largely to dissolved sulphates. On Earth, some microorganisms obtain metabolic energy by reducing sulphate to sulphide — might microbes on Europa do the same thing? If so, they would need a source of electrons, which on Earth is typically provided by organic compounds or dissolved hydrogen.

Zolotov and Shock estimate that hydrothermal reactions between hot rocks and water on the young, warm moon could have generated appreciable amounts of hydrogen, coexisting with dissolved sulphate. This could have supplied the 'fuel' for early life, but it would not have lasted long: although there is plenty of sulphate, the amount of hydrogen is limited. To sustain life for much of Europa's history, other sources of electron donors would have been needed: for example, organic molecules brought by meteorites, recycled 'dead' organic matter in sediments, or recurrent hydrothermal activity that replenished hydrogen or released other reduced volcanic gases.

Philip Ball



Apoptosis

Dead cell talking

Dev. Cell **4**, 587–598 (2003)

Programmed cell death (apoptosis) is a necessity for most living organisms. So too is the rapid removal of the resulting cell corpses by specialized engulfing cells or helpful bystanders, to prevent harmful spillage of cellular constituents. But how are dead cells distinguished from living ones? Swathi Arur *et al.* report that the exposure of annexin I protein on the surface of cell corpses is their way of saying, "Eat me".

Like apoptosis itself, the removal of cell corpses is well coordinated. The engulfing cell binds a dying cell. Its membranes protrude, as if it were wrapping its arms around its target. The tips of the arms fuse and the apoptotic cell is digested.

Arur *et al.* now show that, in dying human cells, annexin I is recruited from the cytosol to discrete patches of the cell membrane. Stuck in the membrane's outer leaflet, it co-localizes with an infamous marker of dying cells, the lipid phosphatidylserine. There it seems to be essential for phosphatidylserine receptors on an engulfing cell to cluster around the dying cell and neatly zip the two cells together.

Marie-Thérèse Heemels

Metallurgy

Bend it, shape it

Science **300**, 464–467 (2003)

A new class of metal alloys developed by Takashi Saito *et al.* combines a remarkable array of unusual and potentially useful properties. The alloys are highly stretchy — both elastically and plastically — and can be made very strong. They also show 'Invar' behaviour, meaning that they expand very little with increasing temperature, and 'Elinvar' behaviour, maintaining constant stiffness over a wide temperature range. These latter properties are important for making delicate engineering components — hairsprings in wrist-watches, for instance.

The new alloys are made of elements from groups IVa and Va of the periodic table — titanium, zirconium, vanadium, niobium and tantalum — along with small amounts of oxygen. They all share three 'magic numbers': an average metal valency of 4.24, a bond order of 2.87, and an electronic *d*-orbital energy of about 2.45 eV. Saito and colleagues say that, unlike most metals, these alloys deform plastically not by local movements of dislocation defects in the crystal structure but by large-scale sliding of defects akin to geological faults. This means that the alloys can undergo large plastic deformation without experiencing 'work hardening', making it possible to shape the cold metals rather like clay.

Philip Ball

The earliest recorded plant virus disease

Pathogenic DNA paints summer foliage gold, and inspired a poet over a millennium ago.

The Man'yōshū, meaning 'collection of ten thousand leaves', is the largest and earliest anthology of Japanese poetry — it contains over 4,500 poems written between the early seventh century and the middle of the eighth century, and provides a glimpse of Japanese life during that period. One poem, attributed to the Empress Kōken and written in the summer of 752 AD (Fig. 1), describes the autumnal appearance of eupatorium plants in summer¹ and is reputedly the earliest written record of the symptoms of a plant virus disease². Here we show that a geminivirus and an accompanying satellite component isolated from affected eupatorium plants are together responsible for the spectacular foliar display that was first noted by the Empress more than a millennium ago.

Eupatorium plants in Japan frequently exhibit a striking yellow pattern on their leaves during summer³, a phenotype that has been correlated with the presence of a geminivirus named eupatorium yellow-vein virus (EpYVV)⁴. Because only a single genomic component had been reported for EpYVV, we suspected that the plants might be further infected with a DNA β-satellite component that is similar to those discovered in

ageratum (a widespread weed) and cotton^{5,6}.

Nucleic acids were extracted from affected plants (*Eupatorium makinoi*, isolate MNS2) originating from the Fukuoka Prefecture in Japan, and a satellite component was amplified from a circular DNA β-satellite template by using the polymerase chain reaction with divergent primers⁷ (clone pGEM-EYVVBM10; GenBank accession number, AJ438938). These primers also amplified naturally occurring recombinants containing geminivirus and satellite sequences, which resembled those previously described⁸, facilitating the isolation of a circular geminivirus DNA component using divergent primers (clone pGEM-EYVVM1; accession number, AJ438936).

Sequence analysis indicated that the geminivirus is an isolate of EpYVV from the Kumamoto Prefecture (88% identity)⁴ and is closely related to Japanese honeysuckle yellow-vein mosaic virus⁹, a geminivirus that also has an accompanying satellite component (our unpublished results). The eupatorium satellite component is distinct from those associated with diseased ageratum and cotton (56% and 53% identity, respectively), although they all contain similar genes and control elements^{5,6}.

To test their infectivity, we constructed cloned tandem repeats of the geminivirus and β-satellite DNA components, and introduced them biolistically⁶ into eupatorium plants. Characteristic yellow-vein symptoms appeared in 5 out of 25 plants (in 2 experiments) from 3–4 weeks post-inoculation (Fig. 2a), and the presence of both components in symptomatic plants was confirmed by Southern-blot analysis (Fig. 2b). The disease was transmitted between eupatorium plants (8 out of 15 plants in 3 experiments) using the whitefly *Bemisia tabaci*⁵, which is responsible for transmitting the disease in the field. These results confirm the aetiology of the disease by fulfilling Koch's postulates¹⁰.

We have shown that the beautiful foliar symptoms in eupatorium plants are caused by a geminivirus–satellite disease complex. To our knowledge, the poem that appears in the Man'yōshū represents the first documented description of a plant virus disease. Although this observation is anecdotal, it implies that such disease complexes were prevalent before the advent of modern intensive agricultural practices, which have encouraged the spread and diversification of geminivirus diseases. Similar disease complexes have now been found in weeds, ornamental plants and economically important crops throughout Africa and

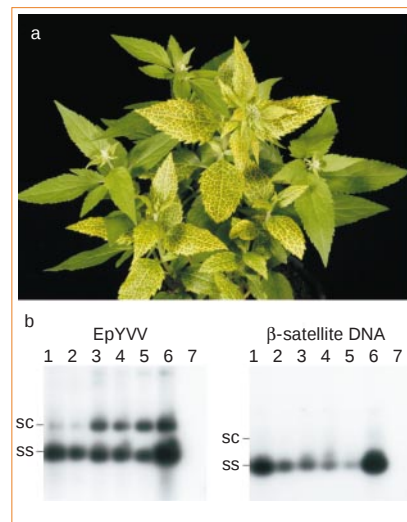


Figure 2 Confirmation of the aetiology of eupatorium yellow-vein disease. **a**, Diseased eupatorium plants exhibiting the yellow-vein phenotype, growing among healthy plants. The disease was transmitted by whiteflies that had acquired the virus from plants infected by geminivirus and satellite cloned components. **b**, Southern-blot analysis of DNA extracted from naturally infected plants (lanes 1 and 2), plants infected using either cloned components (lanes 3 and 4) or viruliferous whiteflies that had acquired viral progeny of the cloned components (lanes 5 and 6), and a healthy plant (lane 7). Blots were hybridized with EpYVV viral DNA (left) and β-satellite DNA (right) probes. SS, single-stranded DNA; SC, supercoiled DNA.

Asia¹¹, indicating that they are diverse and widespread, and represent a serious threat to agriculture in the Old World.

Keith Saunders*, Ian D. Bedford*, Tetsukazu Yahara†, John Stanley*

*Department of Disease and Stress Biology, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

e-mail: john.stanley@bbsrc.ac.uk

†Department of Biology, Kyushu University, 6-10-1 Fukuoka, 812-8581, Japan

1. Suga, T. *The Man'yō-shū: A Complete English Translation in 5–7 Rhythm* (Kanda Inst. Foreign Lang., Tokyo, 1991).
2. Inouye, T. & Osaki, T. *Ann. Phytopath. Soc. Jpn* **46**, 49–50 (1980).
3. Yahara, T. & Oyama, K. *Oecologia* **96**, 310–315 (1993).
4. Onuki, M. & Hanada, K. *J. Gen. Plant Pathol.* **66**, 176–181 (2000).
5. Saunders, K. et al. *Proc. Natl Acad. Sci. USA* **97**, 6890–6895 (2000).
6. Briddon, R. W. et al. *Virology* **285**, 234–243 (2001).
7. Briddon, R. W., Bull, S. E., Mansoor, S., Amin, I. & Markham, P. G. *Mol. Biotechnol.* **20**, 315–318 (2002).
8. Saunders, K., Bedford, I. D. & Stanley, J. *Virology* **282**, 38–47 (2001).
9. Osaki, T., Kobatake, H. & Inouye, T. *Ann. Phytopath. Soc. Jpn* **45**, 62–69 (1979).
10. Koch, R. *Mittheilungen aus dem Kaiserlichen Gesundheitsamt* Vol. 2 (1884).
11. Mansoor, S., Briddon, R. W., Zafar, Y. & Stanley, J. *Trends Plant Sci.* **8**, 128–134 (2003).

Competing financial interests: declared none.

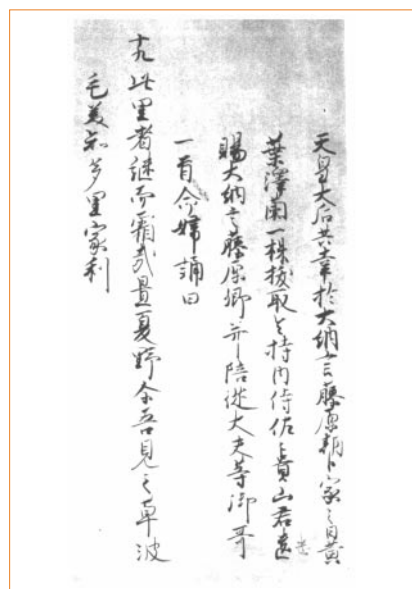


Figure 1 An early handwritten version of a poem by the Empress Kōken describing the autumnal feel of the yellow leaves of eupatorium plants (translated here as grass) growing in summer¹. The text reads: "Perhaps it does frost/ In this village morn by morn/ For the grass I saw in the field of summertime/ Has already turned yellow". The poem, taken from the Rui-shū-ko-shū edition of the Man'yōshū edited by Atsukata Fujiwara (1071–1120), was provided by Nobuyuki Sakamoto and is reproduced by courtesy of the Nara University of Education, Japan.

A new redox-cofactor vitamin for mammals

Nicotinamides and flavins are essential cofactors in enzyme-catalysed reduction–oxidation (redox) reactions and are classified as vitamins because they must be supplied in the diet. Another redox cofactor, pyrroloquinoline quinone (PQQ), was first discovered in bacteria¹ and is also likely to be important in mammals^{2,3}, but the biochemical pathways in which it participates are unknown. Here we identify a PQQ-dependent dehydrogenase enzyme that is crucial for the degradation of the amino acid lysine in mice. PQQ is acting as a mammalian redox cofactor in this reaction, and therefore qualifies as a newcomer to the B group of vitamins.

In animals, lysine is an essential amino acid (that is, one that cannot be made in the body) that is degraded after use to 2-aminoadipic 6-semialdehyde (AAS), which is subsequently oxidized to 2-aminoadipic acid (AAA; Fig. 1a, right arrows)⁴. The molecular features of the AAS-dehydrogenase (AASDH; EC 1.2.1.31) enzyme that catalyses this oxidation of AAS are unknown.

In yeast, lysine is synthesized from AAA by the reverse pathway (Fig. 1a, left arrows)⁵. The reduction of AAA to AAS requires two enzymes that are encoded by the *Lys2* and *Lys5* genes: the *LYS2* protein has NADPH-dependent AAA-reductase activity, whereas *LYS5* catalyses the post-translational 4'-phosphopantetheinylation of *LYS2* (ref. 6). A human homologue of *Lys5* has been cloned⁷, raising the possibility that there is a structural analogue of *LYS2* in animals.

To find the analogue of *LYS2*, we searched the *Drosophila* genome databases for gene product(s) containing functional domains, adenylation and thiolation domains, which serve to capture the substrate in *LYS2* (ref. 6). A BLAST search detected two gene products, EBONY (DDBJ/EMBL/GenBank accession number, CAA11962) and U26 (accession number, AAF52679); EBONY is β -alanyl-dopamine synthetase, but the function of U26 was unknown. In contrast to yeast *LYS2*, which contains a binding domain for the reducing cofactor NADPH, *Drosophila* U26 contains several repeats of the binding motif for PQQ (an oxidizing cofactor); this motif is conserved among bacterial PQQ-dependent dehydrogenases⁸.

To investigate whether U26 could be AAS-dehydrogenase, we cloned mouse U26 complementary DNA (accession number, AB095954), which encodes a protein of 1,100 amino acids that is 23% identical and 43% similar to *Drosophila* U26. Mouse U26 contains one set of adenylation/thiolation

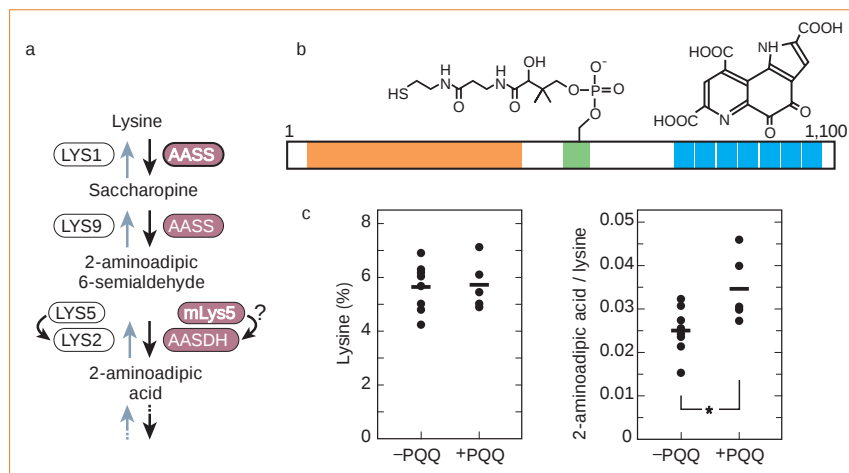


Figure 1 The role of pyrroloquinoline quinone (PQQ) in mammalian lysine degradation. **a**, Metabolic pathways of lysine: right arrows, lysine degradation in mammals; left arrows, lysine biosynthesis in yeast. In yeast, *LYS2* is post-translationally 4'-phosphopantetheinylated in a reaction catalysed by *LYS5*. AAS, AAS-synthetase; AASDH, AAS-dehydrogenase; mLys5, mammalian homologue of *LYS5*. **b**, Arrangement of putative binding domains in murine U26 protein. The adenylation domain (Pfam 00501), thiolation domain (00550), and PQQ-binding motifs (01011) are shown in orange, green and blue, respectively. Serine 592 in the thiolation domain is 4'-phosphopantetheinylated and PQQ (right) is non-covalently bonded by the seven repeats of the PQQ-binding motif near the carboxy terminus. **c**, Peripheral blood concentrations of lysine and 2-aminoadipic acid in PQQ-supplemented and PQQ-deprived mice. Lysine is quantified as a percentage of the total amino acids (left; determined using an amino-acid analyser) and 2-aminoadipic acid is normalized with respect to lysine (right). Horizontal bars, average for each group of mice. Asterisk denotes $P = 0.010$.

domains and seven PQQ-binding motifs (Fig. 1b). The six to eight tandem repeats are also present in bacterial PQQ-dependent dehydrogenases⁸, indicating that mouse U26 could be a PQQ-dependent dehydrogenase.

Northern-blot analysis showed that mouse U26 messenger RNA is expressed in all of the tissues that we tested, with expression being notably high in the heart, liver and kidney (results not shown). In mice fed on a lysine-rich diet (containing 2,000-fold more free lysine than normal), blood concentrations of lysine and AAA were markedly increased (2.4- and 1.5-fold, respectively). The level of U26 mRNA (normalized to that of glyceraldehyde-3-phosphate dehydrogenase) in the liver and heart of lysine-loaded mice was significantly higher (1.6- and 1.4-fold, respectively; $P = 0.008$ for each) than in control mice. Mouse AAS-synthetase (AAS in Fig. 1a) and *Lys5* transcripts were also increased in the lysine-loaded mice (results not shown). These results indicate that U26, as well as AAS-synthetase and *LYS5*, is involved in the lysine-degradation pathway in mice.

To test the importance of PQQ in AAS-dehydrogenase activity, we examined the effects of a PQQ-deficient diet on mice. After being fed on a chemically defined diet^{2,3} that was either devoid of PQQ or supplemented with PQQ (1,000 ng PQQ/2Na per g of food), we measured the blood levels of lysine and AAA in the two groups of mice. AAA concentrations were significantly decreased in PQQ-deprived mice, whereas lysine levels remained the same (Fig. 1c); the intermediates saccharo-

pine (Fig. 1a) and AAS were not detected. This decrease in AAA may be indicative of impaired AAS-dehydrogenase (U26) activity resulting from a PQQ deficiency in deprived mice, pointing to a requirement for PQQ as a redox cofactor for the proper functioning of AAS-dehydrogenase.

PQQ is found in various foods, including vegetables and meat^{9,10}. When mice are fed a PQQ-deficient diet, they grow slowly, have fragile skin and a reduced immune response, and do not reproduce well^{2,3}. On the basis of our demonstration of its molecular function, we propose that PQQ should be classified as a new B vitamin, joining niacin/nicotinic acid (vitamin B3) and riboflavin (vitamin B2), the redox-cofactor derivatives of which are $\text{NAD}^+/\text{NADP}^+$ and FAD/FMN, respectively.

Takaoki Kasahara, Tadaaki Kato

Laboratory for Molecular Dynamics of Mental Disorders, Brain Science Institute, RIKEN, Wako-shi, Saitama 351-0198, Japan
e-mail: kato@brain.riken.go.jp

- Salisbury, S. A., Forrest, H. S., Cruse, W. B. & Kennard, O. *Nature* **280**, 843–844 (1979).
- Killgore, J. et al. *Science* **245**, 850–852 (1989).
- Steinberg, F. M., Gershwin, M. E. & Rucker, R. B. *J. Nutr.* **124**, 744–753 (1994).
- Higashino, K., Fujioka, M. & Yamamura, Y. *Arch. Biochem. Biophys.* **142**, 606–614 (1971).
- Bhattacharjee, J. K. *Crit. Rev. Microbiol.* **12**, 131–151 (1985).
- Ehmann, D. E., Gehring, A. M. & Walsh, C. T. *Biochemistry* **38**, 6171–6177 (1999).
- Praphanphoj, V., Sacksteder, K. A., Gould, S. J., Thomas, G. H. & Geraghty, M. T. *Mol. Genet. Metab.* **72**, 336–342 (2001).
- Mathews, F. S. *Methods Enzymol.* **258**, 191–216 (1995).
- Flückiger, R., Woodtli, T. & Gallop, P. M. *Biochem. Biophys. Res. Commun.* **153**, 353–358 (1988).
- Kumazawa, T., Seno, H., Urakami, T., Matsumoto, T. & Suzuki, O. *Biochim. Biophys. Acta* **1156**, 62–66 (1992).

Competing financial interests: declared (see online version).

Why do birds engage in extra-pair copulation?

Female birds often copulate with males that are not their social partner, but why they do this is unclear¹. Blomqvist *et al.*² present evidence from three wader species showing that extra-pair fertilizations are more likely when social mates are genetically closely related. We argue, however, that their conclusions are undermined on technical and theoretical grounds, and that there is still no convincing evidence that extra-pair copulation in birds is influenced by the relatedness between partners.

Blomqvist *et al.* used a single multilocus (minisatellite) DNA-fingerprinting probe to estimate "genetic similarity". Although multilocus probes will detect extra-pair paternity¹ and maternity³ (brood parasitism), band-sharing coefficients (calculated by comparing bands on DNA fingerprints) are unreliable for determining the degree of genetic similarity between two individuals⁴. Blomqvist *et al.* present data that illustrate this problem, reporting that band-sharing coefficients between individual parents and their offspring vary between 0.31 and 0.79, when in fact their genetic similarity is always the same, as each parent passes half of its autosomal alleles to each offspring.

Multilocus DNA fingerprints have been used successfully to compare the mean genetic similarity of two samples^{3,5}, but the results of Blomqvist *et al.* reveal two patterns — quasi-parasitism and kin recognition — that should be rare or absent in birds in general, and in waders in particular. They find evidence of quasi-parasitism in four of six extra-pair matings in common sandpipers (*Actitis hypoleucos*) and Kentish plovers (*Charadrius alexandrinus*), although not in western sandpipers (*Calidris mauri*).

Quasi-parasitism occurs when an extra-pair female with whom a male has copulated visits the male's nest and lays an egg. Of 130 bird species for which molecular analysis for parentage has been carried out⁶, only one (the sand martin, *Riparia riparia*)³ shows evidence of quasi-parasitism that is supported by behavioural observations. This is unsurprising, as *R. riparia* breeds in dense colonies where females nest only a few centimetres apart, offering opportunities for both quasi-parasitism and extra-pair paternity. The waders studied by Blomqvist *et al.*, however, are territorial, with neighbours nesting tens to hundreds of metres apart, and are socially and (largely) sexually monogamous. Also, a larger study of common sandpipers using two multilocus probes and behavioural observations found no evidence of quasi-parasitism⁶.

As these waders lay no more than four eggs per clutch, and Blomqvist *et al.* found no evidence of supernumerary eggs in the clutches that they studied, pair males or extra-pair females must have removed an intact egg (each weighing $\geq 20\%$ of adult body mass) to accommodate one from the extra-pair female. This would be an extraordinary physical feat for these birds and has not previously been reported in waders. The removed within-pair egg would almost certainly have been fertilized by the male, unlike that laid by the extra-pair female, for which the likelihood of the male's paternity would be lower.

Blomqvist *et al.* conclude that a pair member that detects a strong genetic similarity with its partner seeks an extra-pair mating as an adaptive tactic to avoid the costs of inbreeding. Although they do not estimate coefficients of relatedness (*r*) from their DNA fingerprints, their Fig. 1 indicates (using methods in ref. 4) that individual birds would have to have recognized their partner as third- or higher-order kin ($r \leq 0.125$) in seven out of eight pairs with quasi-parasitism or extra-pair paternity. Having established this, the birds would then copulate with more distantly related individuals. However, the only (limited) evidence so far for kin recognition in birds involves first- and possibly second-order relatives, such as peacocks that recognize their brothers and half-brothers⁷.

Despite the potentially serious deleterious effects of inbreeding in wild bird populations, there is little to suggest that birds use kin discrimination to avoid inbreeding during primary mate selection⁸. Moreover, kin discrimination is unlikely in migratory birds such as waders that have low natal philopatry, as relatives rarely encounter one another.

Blomqvist *et al.*² should have used microsatellite markers for their molecular analysis, which are more reliable for estimating coefficients of relatedness between individuals⁹, to be certain of the genetic relationship between mates and the incidence of quasi-parasitism in these waders. Even then, alternative interpretations are possible. Meanwhile, extra-pair mating in birds remains to be explained.

Simon C. Griffith*, Robert Montgomerie†

*Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
e-mail: simon.griffith@zoology.oxford.ac.uk
†Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada

1. Griffith, S. C., Owens, I. P. F. & Thuman, K. A. *Mol. Ecol.* **11**, 2195–2212 (2002).
2. Blomqvist, D. *et al.* *Nature* **419**, 613–615 (2002).
3. Alves, M. A. S. & Bryant, D. M. *Anim. Behav.* **56**, 1323–1331 (1998).
4. Reeve, H. K., Westneat, D. F. & Queller, D. C. *Mol. Ecol.* **1**, 223–232 (1992).
5. Bensch, S., Hasselquist, D. & von Schantz, T. *Evolution* **48**, 317–326 (1994).

6. Mee, A., Whitfield, D. P., Thompson, D. B. A. & Burke, T. *Anim. Behav.* (in the press).
7. Petrie, M., Krupa, A. & Burke, T. *Nature* **401**, 155–157 (1999).
8. Keller, L. F. & Arcese, P. *Am. Nat.* **152**, 380–392 (1998).
9. Queller, D. C., Strassmann, J. E. & Hughes, C. R. *Trends Ecol. Evol.* **8**, 285–288 (1993).

Blomqvist et al. reply — Griffith and Montgomerie question our finding¹ that extra-pair parentage in shorebirds correlates with the genetic similarity of social mates, on the grounds that band-sharing coefficients (calculated by comparing bands on DNA fingerprints) are unreliable for determining the degree of genetic similarity between two individuals.

Although band-sharing does not give an exact measure of relatedness between two individuals, it does provide an index that reflects their relatedness^{2,3}. Band-sharing coefficients can therefore be used justifiably to test a null hypothesis of no difference in average relatedness between groups^{1,4,5}. Random variation in the scores from individual pairs reduces the chances that the null hypothesis will be refuted, and is therefore unlikely to explain our significant results. It is also difficult to see how our analysis could create spurious significant differences in band-sharing between pairs with and without extra-pair parentage, in each of three different species.

Griffith and Montgomerie also doubt our proposed cases of quasi-parasitism (extra-pair maternity), but there is genetic evidence for quasi-parasitism in at least five other bird species, some of which are territorial^{6–10}. We recorded quasi-parasitism in 2 out of 15 broods¹ of the common sandpiper (*Actitis hypoleucos*), which was not detected in the study cited by Griffith and Montgomerie. This is not surprising, however, as in some species the frequency of extra-pair paternity also varies between populations¹¹.

Griffith and Montgomerie argue that egg removal by the parasitic female or the pair male is highly improbable. However, shorebirds (waders) frequently remove damaged eggs from the nest¹², and adults of some species will roll intact eggs out of the nest to a distance of many metres¹³. Removal of an egg is not necessary to explain our results, however. As we pointed out¹, the host female may have stopped producing eggs because the parasitic female laid her egg(s) early in the laying sequence of the host female. This is not unlikely, given that laying intervals are variable in shorebirds. For example, female Kentish plovers (*Charadrius alexandrinus*) may wait up to five days between laying consecutive eggs (our unpublished observations).

We do not agree with Griffith and Montgomerie that we should have used microsatellite markers, as these would be unlikely to yield qualitatively different results. Band-sharing coefficients from

multilocus DNA fingerprinting, based on a single probe, have been shown to reflect genetic relatedness accurately in the red-winged blackbird (*Agelaius phoeniceus*)³. We maintain that our application of this method in testing for differences in the average level of relatedness between groups is justified.

We agree with Griffith and Montgomerie that alternative explanations may need to be considered before concluding that mate choice is based on genetic diversity and kin discrimination. But, as we saw a similar pattern in three species with different ecology and social behaviour¹, we contend that our proposal of adaptive extra-pair copulation with genetically dissimilar mates warrants further testing in these and other species.

Donald Blomqvist, Malte Andersson, Clemens Küpper, Innes C. Cuthill, János Kis, Richard B. Lanctot, Brett K. Sandercock, Tamás Székely, Johan Wallander, Bart Kempnaers*

*Max Planck Research Centre for Ornithology, PO Box 1564, 82305 Starnberg (Seewiesen), Germany
e-mail: b.kempnaers@erl.ornithol.mpg.de

1. Blomqvist, D. et al. *Nature* **419**, 613–615 (2002).
2. Lynch, M. *DNA Fingerprinting: Approaches and Applications* (eds Burke, T. et al.) 113–126 (Birkhäuser, Basel, 1991).
3. Reeve, H. K., Westneat, D. F. & Queller, D. C. *Mol. Ecol.* **1**, 223–232 (1992).
4. Brown, K. M. *Anim. Behav.* **56**, 1529–1543 (1998).
5. Petrie, M., Krupa, A. & Burke, T. *Nature* **401**, 155–157 (1999).
6. Alves, M. A. S. & Bryant, D. M. *Anim. Behav.* **56**, 1323–1331 (1998).
7. Hoi, H. & Hoi-Leitner, M. *Behav. Ecol.* **8**, 113–119 (1997).
8. Reyer, H. U., Bollmann, K., Schläpfer, A. R., Schymainda, A. & Kleck, G. *Behav. Ecol.* **8**, 534–543 (1997).
9. Schleicher, B., Hoi, H., Valera, F. & Hoi-Leitner, M. *Behaviour* **134**, 941–959 (1997).
10. Václav, K., Hoi, H. & Blomqvist, D. *Behav. Ecol.* (in the press).
11. Petrie, M. & Kempnaers, B. *Trends Ecol. Evol.* **13**, 52–58 (1998).
12. Sandercock, B. K. *Oecologia* **110**, 50–59 (1997).
13. Chylarecki, P. *Wader Study Group Bull.* **94**, 21–22 (2001).

COMMUNICATIONS ARISING

Astrophysics

How did the metals in a giant star originate?

The chemical composition of stars with extremely low metal contents (taking ‘metals’ to mean all elements other than hydrogen and helium) provides us with information on the masses of the stars that produced the first metals. Such a direct connection is not possible, however, if the surface of the star has been polluted by enriched material, either dredged from the star’s interior or transferred from a companion star. Here we argue that, in the case of HE0107–5240 (ref. 1), the most iron-poor star known, the oxygen abundance could be a discriminant: a ratio of [O/Fe] exceeding +3.5 would favour a pristine origin of metals, whereas an [O/Fe] ratio of less than +3 would favour the pollution

hypothesis. Using this criterion, we suggest how the required information on oxygen abundance might be obtained.

HE0107–5240 shows carbon, nitrogen and sodium enhanced, respectively, by factors of 10^4 , $10^{2.3}$ and 10 relative to iron, whereas magnesium, which is usually more abundant in metal-poor stars, is present in almost the same amounts as iron; no single supernova (SNII) model seems to show this pattern². Although several models may be worth considering, we evaluate two here: a pristine origin due to the combined enrichment of at least two SNIIs, or a pollution of the surface of the star after its formation.

If the metals are pristine, then HE0107–5240 must have formed from a cloud enriched by the ejecta of at least two zero-metallicity SNIIs of quite different initial masses. The first supernova, which was presumably rather massive, would have produced the light elements observed in HE0107–5240, but none of the heavier ones, because of an extensive fallback, up to the base of the helium shell. This would have prevented the ratios [C/N] and [Mg/C] from becoming too high. The second supernova, which was less massive, would have provided all of the elements heavier than and including magnesium.

In contrast, in the case of pollution enrichment, HE0107–5240 would be either a low-mass, zero-metallicity star that accreted matter enriched by the first generation of SNII, or a second-generation low-mass star of very low metallicity. The high [C,N,Na/Fe] ratios would be the result of subsequent enrichment of the surface of the star, due either to an internal process or to the accretion of matter synthesized by an asymptotic giant branch (AGB) star in a binary system.

One way to discriminate between pristine and pollution origins of C and N is to measure the abundance of oxygen. In the pristine case, the more massive supernova would provide a large amount of oxygen, leading to a high [O/Fe] ($> +3.5$), as can be derived from the yields of zero-metal massive stars². In the pollution case, a lower [O/Fe] is expected ($< +3.0$), as implied by the yields of zero-metal, intermediate-mass stars³.

An upper limit of 0.1 pm (where 1 pm is 10^{-12} m) on the equivalent width of the [OI] 630-nm line would provide a limit on oxygen, namely [O/H] < -2.3 or [O/Fe] $< +3$. With the Ultraviolet–Visual Echelle Spectrograph on the European Southern Observatory’s Very Large Telescope, we estimate that such a limit could be achieved in about 30 h of exposure. The situation is slightly better for OH ultraviolet lines, for which we estimate that an upper limit of [O/Fe] $< +2.0$ could be derived even at moderate signal-to-noise ratios (> 20), which should be achieved in about

4 h of exposure. At this signal-to-noise ratio, it is possible to discriminate between [O/Fe] $< +3.0$ and [O/Fe] $> +3.5$; however, lower oxygen abundances would be very difficult to measure, whatever the signal-to-noise ratio. As high-resolution spectra have already been obtained, including the ultraviolet region containing the OH lines, for a total of about 30 h of exposure (ESO/ST-ECF Science Archive), it may now be possible to carry out the test that we propose.

Among extremely metal-poor stars, oxygen has been observed only in the star CS 22949–037 (ref. 4), for which [O/Fe] = +2.0; all other known stars have oxygen features that are below the detection threshold. However, CS 29498–043 (ref. 5) has a very similar pattern of marked overabundance of the lighter elements; its oxygen abundance, which has not yet been measured, may prove to be strongly enhanced as well.

We note that the high observed [C/N] in HE0107–5240 is difficult to explain in terms of pollution, whether internal or from an AGB companion, because of the high abundance of primary nitrogen in both cases. We argue that a pristine origin is more likely, which implies that the [O/Fe] ratio is higher than in any other known metal-poor star.

Piercarlo Bonifacio*, Marco Limongi†, Alessandro Chieffi‡

*Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Trieste, 34131 Trieste, Italy
e-mail: bonifaci@ts.astro.it

†Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Roma, 00040 Monteporzio Catone, Roma, Italy

‡Istituto di Astrofisica Spaziale, CNR, 00133 Roma, Italy

1. Christlieb, N. et al. *Nature* **419**, 904–906 (2002).
2. Chieffi, A. & Limongi, M. *Astrophys. J.* **577**, 281–294 (2002).
3. Chieffi, A., Limongi, M., Dominguez, I. & Straniero, O. in *New Questions in Stellar Astrophysics: The Link Between Stars and Cosmology* (eds Chavez, M. et al.) 47–54 (Kluwer Academic, Dordrecht, The Netherlands, 2002).
4. Depagne, E. et al. *Astron. Astrophys.* **390**, 187–198 (2002).
5. Aoki, W., Norris, J. E., Ryan, S. G., Beers, T. C. & Ando, H. *Astrophys. J.* **576**, L141–L144 (2002).

addendum

Adult persistence of head-turning asymmetry

O. Güntürkün

Nature **421**, 711 (2003)

Some confusion over the definition of the term ‘right kiss’ as used in this communication has resulted from the depiction of Auguste Rodin’s statue *The Kiss*, in which the male turns his head in the vertical plane to the left. However, this sculpture only served to illustrate the definition of a right kiss, which corresponds to a position in which the nose of each participant is to the right of the nose of the other. The strict criteria applied during my observations excluded couples in the position of Rodin’s statue and ensured that a right kiss incorporated a turn and a tilt with the head to the right side.

A vision for the future of genomics research

A blueprint for the genomic era.

Francis S. Collins, Eric D. Green, Alan E. Guttmacher and Mark S. Guyer on behalf of the US National Human Genome Research Institute*

The completion of a high-quality, comprehensive sequence of the human genome, in this fiftieth anniversary year of the discovery of the double-helical structure of DNA, is a landmark event. The genomic era is now a reality.

In contemplating a vision for the future of genomics research, it is appropriate to consider the remarkable path that has brought us here. The rollfold (Figure 1) shows a timeline of landmark accomplishments in genetics and genomics, beginning with Gregor Mendel's discovery of the laws of heredity¹ and their rediscovery in the early days of the twentieth century. Recognition of DNA as the hereditary material², determination of its structure³, elucidation of the genetic code⁴, development of recombinant DNA technologies^{5,6}, and establishment of increasingly automatable methods for DNA sequencing^{7–10} set the stage for the Human Genome Project (HGP) to begin in 1990 (see also www.nature.com/nature/DNA50). Thanks to the vision of the original planners, and the creativity and determination of a legion of talented scientists who decided to make this project their overarching focus, all of the initial objectives of the HGP have now been achieved at least two years ahead of expectation, and a revolution in biological research has begun.

The project's new research strategies and experimental technologies have generated a steady stream of ever-larger and more complex genomic data sets that have poured into public databases and have transformed the study of virtually all life processes. The genomic approach of technology development and large-scale generation of community resource data sets has introduced an important new dimension into biological and biomedical research. Interwoven advances in genetics, comparative genomics, high-throughput biochemistry and bioinformatics

are providing biologists with a markedly improved repertoire of research tools that will allow the functioning of organisms in health and disease to be analysed and comprehended at an unprecedented level of molecular detail. Genome sequences, the bounded sets of information that guide biological development and function, lie at the heart of this revolution. In short, genomics has become a central and cohesive discipline of biomedical research.

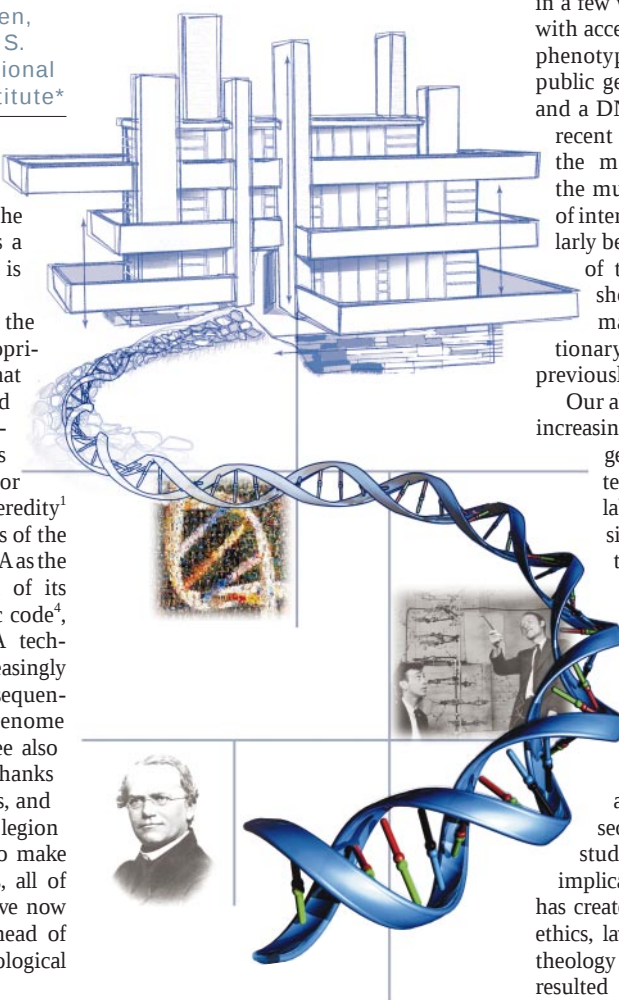
The practical consequences of the emergence of this new field are widely apparent. Identification of the genes responsible for human mendelian diseases, once a herculean task requiring large research teams, many years of hard work, and an uncertain outcome, can now be routinely accomplished

in a few weeks by a single graduate student with access to DNA samples and associated phenotypes, an Internet connection to the public genome databases, a thermal cycler and a DNA-sequencing machine. With the recent publication of a draft sequence of the mouse genome¹¹, identification of the mutations underlying a vast number of interesting mouse phenotypes has similarly been greatly simplified. Comparison of the human and mouse sequences shows that the proportion of the mammalian genome under evolutionary selection is more than twice that previously assumed.

Our ability to explore genome function is increasing in specificity as each subsequent genome is sequenced. Microarray technologies have catapulted many laboratories from studying the expression of one or two genes in a month to studying the expression of tens of thousands of genes in a single afternoon¹². Clinical opportunities for gene-based pre-symptomatic prediction of illness and adverse drug response are emerging at a rapid pace, and the therapeutic promise of genomics has ushered in an exciting phase of expansion and exploration in the commercial sector¹³. The investment of the HGP in studying the ethical, legal and social implications of these scientific advances has created a talented cohort of scholars in ethics, law, social science, clinical research, theology and public policy, and has already resulted in substantial increases in public awareness and the introduction of significant (but still incomplete) protections against misuses such as genetic discrimination (see www.genome.gov/PolicyEthics).

These accomplishments fulfil the expansive vision articulated in the 1988 report of the National Research Council, *Mapping and Sequencing the Human Genome*¹⁴. The successful completion of the HGP this year thus represents an opportunity to look forward and offer a blueprint for the future of genomics research over the next several years.

The vision presented here addresses a different world from that reflected in earlier plans published in 1990, 1993 and 1998 (refs 15–17). Those documents addressed the goals of the 1988 report, defining detailed paths towards the development of genome-



*Endorsed by the National Advisory Council for Human Genome Research, whose members are Vickie Yates Brown, David R. Burgess, Wylie Burke, Ronald W. Davis, William M. Gelbart, Eric T. Juengst, Bronya J. Keats, Raju Kucherlapati, Richard P. Lifton, Kim J. Nickerson, Maynard V. Olson, Janet D. Rowley, Robert Tepper, Robert H. Waterston and Tadataka Yamada.

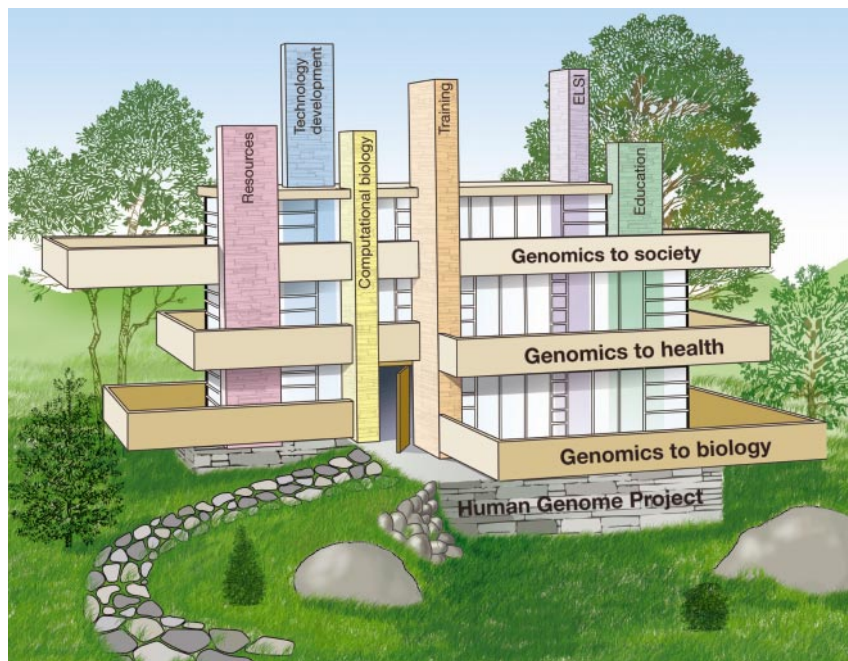


Fig 2 The future of genomics rests on the foundation of the Human Genome Project.

analysis technologies, the physical and genetic mapping of genomes, and the sequencing of model organism genomes and, ultimately, the human genome. Now, with the effective completion of these goals, we offer a broader and still more ambitious vision, appropriate for the true dawning of the genomic era. The challenge is to capitalize on the immense potential of the HGP to improve human health and well-being.

The articulation of a new vision is an opportunity to explore transformative new approaches to achieve health benefits. Although genome-based analysis methods are rapidly permeating biomedical research, the challenge of establishing robust paths

from genomic information to improved human health remains immense. Current efforts to meet this challenge are largely organized around the study of specific diseases, as exemplified by the missions of the disease-oriented institutes at the US National Institutes of Health (NIH, www.nih.gov) and numerous national and international governmental and charitable organizations that support medical research. The National Human Genome Research Institute (NHGRI), in budget terms a rather small (less than 2%) component of the NIH, will work closely with all these organizations in exploring and supporting these biomedical research capabilities. In addition, we envi-

sion a more direct role for both the extramural and intramural programmes of the NHGRI in bringing a genomic approach to the translation of genomic sequence information into health benefits.

The NHGRI brings two unique assets to this challenge. First, it has close ties to a scientific community whose direct role over the past 13 years in bringing about the genomic revolution provides great familiarity with its potential to transform biomedical research. Second, the NHGRI's long-standing mission, to investigate the broadest possible implications of genomics, allows unique flexibility to explore the whole spectrum of human health and disease from the fresh perspective of genome science. By engaging the energetic and interdisciplinary genomics-research community more directly in health-related research and by exploiting the NHGRI's ability to pursue opportunities across all areas of human biology, the institute seeks to participate directly in translating the promises of the HGP into improved human health.

To fully achieve this goal, the NHGRI must also continue in its vigorous support of another of its vital missions — the coupling of its scientific research programme with research into the social consequences of increased availability of new genetic technologies and information. Translating the success of the HGP into medical advances intensifies the need for proactive efforts to ensure that benefits are maximized and harms minimized in the many dimensions of human experience.

A reader's guide

The vision for genomics research detailed here is the outcome of almost two years of intense discussions with hundreds of scientists and members of the public, in more than a dozen workshops and numerous individual consultations (see www.genome.gov/About/Planning). The vision is formulated into three major themes — genomics to biology, genomics to health, and genomics to society — and six crosscutting elements.

We envisage the themes as three floors of a building, firmly resting on the foundation of the HGP (Figure 2). For each theme, we present a series of grand challenges, in the spirit of the proposals put forward for mathematics by David Hilbert at the turn of the twentieth century¹⁸. These grand challenges are intended to be bold, ambitious research targets for the scientific community. Some can be planned on specific timescales, others are less amenable to that level of precision. We list the grand challenges in an order that makes logical sense, not representing priority. The challenges are broad in sweep, not parochial — some can be led by the NHGRI alone, whereas others will be best pursued in partnership with other organizations. Below, we clarify areas in which the NHGRI intends to play a leading role.

BOX 1 Resources



One of the key and distinctive objectives of the Human Genome Project (HGP) has been the generation of large, publicly available, comprehensive sets of reagents and data (scientific resources or 'infrastructure') that, along with other new, powerful technologies, comprise a toolkit for genomics-based research. Genomic maps and sequences are the most obvious examples. Others include databases of sequence variation, clone libraries and collections of anonymous cell lines. The continued generation of such resources is critical, in particular:

- ◆ Genome sequences of key mammals, vertebrates, chordates, and invertebrates
- ◆ Comprehensive reference sets of coding sequences from key species in various formats, for example, full-length cDNA sequences and corresponding clones, oligonucleotide primers, and microarrays

- ◆ Comprehensive collections of knockouts and knock-downs of all genes in selected animals to accelerate the development of models of disease
- ◆ Comprehensive reference sets of proteins from key species in various formats, for example in expression vectors, with affinity tags and spotted onto protein chips
- ◆ Comprehensive sets of protein affinity reagents
- ◆ Databases that integrate sequences with curated information and other large data sets, as well as tools for effective mining of the data
- ◆ Cohort populations for studies designed to identify genetic contributors to health and to assess the effect of individual gene variants on disease risk, including a 'healthy' cohort
- ◆ Large libraries of small molecules, together with robotic methods to screen them and access to medicinal chemistry for follow-up, to provide investigators easy and affordable access to these tools

The six critically important crosscutting elements are relevant to all three thematic areas. They are: resources (Box 1); technology development (Box 2); computational biology (Box 3); training (Box 4); ethical, legal and social implications (ELSI, Box 5); and education (Box 6). We also stress the critical importance of early, unfettered access to genomic data in achieving maximum public benefit. Finally, we propose a series of 'quantum leaps', achievements that would lead to substantial advances in genomics research and its applications to medicine. Some of these may seem overly bold, but no laws of physics need to be violated to achieve them. Such leaps would have profound implications, just as the dreams of the mid-1980s about the complete sequence of the human genome have been realized in the accomplishments now being celebrated.

I Genomics to biology

Elucidating the structure and function of genomes

The broadly available genome sequences of human and a select set of additional organisms represent foundational information for biology and biomedicine. Embedded within this as-yet poorly understood code are the genetic instructions for the entire repertoire of cellular components, knowledge of which is needed to unravel the complexities of biological systems. Elucidating the structure of genomes and identifying the function of the myriad encoded elements will allow connections to be made between genomics and biology and will, in turn, accelerate the exploration of all realms of the biological sciences.

For this, new conceptual and technological approaches will be needed to:

- ◆ Develop a comprehensive and comprehensible catalogue of all of the components encoded in the human genome.
- ◆ Determine how the genome-encoded components function in an integrated manner to perform cellular and organismal functions.
- ◆ Understand how genomes change and take on new functional roles.

Grand Challenge I-1 Comprehensively identify the structural and functional components encoded in the human genome

Although DNA is relatively simple and well understood chemically, the human genome's structure is extraordinarily complex and its function is poorly understood. Only 1–2% of its bases encode proteins⁷, and the full complement of protein-coding sequences still remains to be established. A roughly equivalent amount of the non-coding portion of the genome is under active selection¹¹, suggesting that it is also functionally important, yet vanishingly little is known about it. It



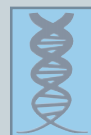
probably contains the bulk of the regulatory information controlling the expression of the approximately 30,000 protein-coding genes, and myriad other functional elements, such as non-protein-coding genes and the sequence determinants of chromosome dynamics. Even less is known about the function of the roughly half of the genome that consists of highly repetitive sequences or of the remaining non-coding, non-repetitive DNA.

The next phase of genomics is to catalogue, characterize and comprehend the entire set of functional elements encoded in the human and other genomes. Compiling this genome 'parts list' will be an immense challenge. Well-known classes of functional

elements, such as protein-coding sequences, still cannot be accurately predicted from sequence information alone. Other types of known functional sequences, such as genetic regulatory elements, are even less well understood; undoubtedly new types remain to be defined, so we must be ready to investigate novel, perhaps unexpected, ways in which DNA sequence can confer function. Similarly, a better understanding of epigenetic changes (for example, methylation and chromatin remodelling) is needed to comprehend the full repertoire of ways in which DNA can encode information.

Comparison of genome sequences from evolutionarily diverse species has emerged as a powerful tool for identifying functionally important genomic elements. Initial analyses of available vertebrate genome sequences^{7,11,19} have revealed many previously undiscovered protein-coding sequences. Mammal-to-mammal sequence comparisons have revealed large numbers of homologies in non-coding regions¹¹, few of which can be defined in functional terms. Further comparisons of sequences derived from multiple species, especially those occupying distinct evolutionary positions, will lead to significant refinements in our understanding of the functional importance of conserved sequences²⁰. Thus, the generation of additional genome sequences from several well-chosen species is crucial to the functional characterization of the human genome (Box 1). The generation of such large sequence data sets will benefit from further advances in sequencing technology that yield significant cost reductions (Box 2). The study of sequence variation within species will also be important in defining the functional nature of some sequences (see Grand Challenge I-3).

BOX 2 Technology development



The Human Genome Project was aided by several 'breakthrough' technological developments, including Sanger DNA sequencing and its automation, DNA-based genetic

markers, large-insert cloning systems and the polymerase chain reaction. During the project, these methods were scaled up and made more efficient by 'evolutionary' advances, such as automation and miniaturization. New technologies, including capillary-based sequencing and methods for genotyping single-nucleotide polymorphisms, have recently been introduced, leading to further improvements in capacity for genomic analyses. Even newer approaches, such as nanotechnology and microfluidics, are being developed, and hold great promise, but further advances are still needed. Some examples are:

- ◆ Sequencing and genotyping technologies to reduce costs further and increase access to a wider range of investigators
- ◆ Identification and validation of functional

elements that do not encode protein

- ◆ *In vivo*, real-time monitoring of gene expression and the localization, specificity, modification and activity/kinetics of gene products in all relevant cell types
- ◆ Modulation of expression of all gene products using, for example, large-scale mutagenesis, small-molecule inhibitors and knock-down approaches (such as RNA-mediated inhibition)
- ◆ Monitoring of the absolute abundance of any protein (including membrane proteins, proteins at low abundance and all modified forms) in any cell
- ◆ Improved imaging methods that allow non-invasive molecular phenotyping
- ◆ Correlating genetic variation to human health and disease using haplotype information or comprehensive variation information
- ◆ Laboratory-based phenotyping, including the use of protein affinity reagents, proteomic approaches and analysis of gene expression
- ◆ Linking molecular profiles to biology, particularly pathway biology to disease

Effective identification and analysis of functional genomic elements will require increasingly powerful computational capabilities, including new approaches for tackling ever-growing and increasingly complex data sets and a suitably robust computational infrastructure for housing, accessing and analysing those data sets (Box 3). In parallel, investigators will need to become increasingly adept in dealing with this treasure trove of new information (Box 4). As a better understanding of genome function is gained, refined computational tools for *de novo* prediction of the identity and behaviour of functional elements should emerge²¹.

Complementing the computational detection of functional elements will be the generation of additional experimental data by high-throughput methodologies. One example is the production of full-length complementary DNA (cDNA) sequences (see, for example, mgc.nci.nih.gov and www.fruitfly.org/EST/full.shtml). Major challenges inherent in programmes to discover genes are the experimental identification and validation of alternate splice forms and messenger RNAs expressed in a highly restricted fashion. Even more challenging is the experimental validation of functional elements that do not encode protein (for example, regulatory regions and non-coding RNA sequences). High-throughput approaches to identify them (Box 2) will be needed to generate the experimental data that will be necessary to develop, confirm and enhance computational methods for detecting functional elements in genomes.

Because current technologies cannot yet identify all functional elements, there is a need for a phased approach in which new methodologies are developed, tested on a pilot scale and finally applied to the



entire human genome. Along these lines, the NHGRI recently launched the Encyclopedia of DNA Elements (ENCODE) Project (www.genome.gov/Pages/Research/ENCODE) to identify all the functional elements in the human genome. In a pilot project, systematic strategies for identifying all functionally important genomic elements will be developed and tested using a selected 1% of the human genome. Parallel projects involving well-studied model organisms, for example, yeast, nematode and fruitfly, are ongoing. The lessons learned will serve as the basis for implementing a broader programme for the entire human genome.

Grand Challenge I-2 Elucidate the organization of genetic networks and protein pathways and establish how they

contribute to cellular and organismal phenotypes

Genes and gene products do not function independently, but participate in complex, interconnected pathways, networks and molecular systems that, taken together, give rise to the workings of cells, tissues, organs and organisms. Defining these systems and determining their properties and interactions is crucial to understanding how biological systems function. Yet these systems are far more complex than any problem that molecular biology, genetics or genomics has yet approached. On the basis of previous experience, one effective path will begin with the study of relatively simple model organisms²², such as bacteria and yeast, and then extend the early findings to more complex organisms, such as mouse and human. Alternatively, focusing on a few well-characterized systems in mammals will be a useful test of the approach (see, for example, www.signaling-gateway.org).

Understanding biological pathways, networks and molecular systems will require information from several levels. At the genetic level, the architecture of regulatory interactions will need to be identified in different cell types, requiring, among other things, methods for simultaneously monitoring the expression of all genes in a cell¹². At the gene-product level, similar techniques that allow *in vivo*, real-time measurement of protein expression, localization, modification and activity/kinetics will be needed (Box 2). It will be important to develop, refine and scale up techniques that modulate gene expression, such as conventional gene-knockout methods²³, newer knock-down approaches²⁴ and small-molecule inhibitors²⁵ to establish the temporal and cellular expression pattern of individual proteins and to determine the functions of those proteins. This is a key first step towards assigning all genes and their products to functional pathways.

The ability to monitor all proteins in a cell simultaneously would profoundly improve our ability to understand protein pathways and systems biology. A critical step towards gaining a complete understanding of systems biology will be to take an accurate census of the proteins present in particular cell types under different physiological conditions. This is becoming possible in some model systems, such as microorganisms²⁶. It will be a major challenge to catalogue proteins present in low abundance or in membranes. Determining the absolute abundance of each protein, including all modified forms, will be an important next step. A complete interaction map of the proteins in a cell, and their cellular locations, will serve as an atlas for the biological and medical explorations of cellular metabolism²⁷ (see www.nrcam.uchc.edu, for example). These and other related areas constitute the developing field of proteomics.

BOX 3 Computational biology



Computational methods have become intrinsic to modern biological research, and their importance can only increase as large-scale methods for data generation become more prominent, as the amount and complexity of the data increase, and as the questions being addressed become more sophisticated. All future biomedical research will integrate computational and experimental components. New computational capabilities will enable the generation of hypotheses and stimulate the development of experimental approaches to test them. The resulting experimental data will, in turn, be used to generate more refined models that will improve overall understanding and increase opportunities for application to disease. The areas of computational biology critical to the future of genomics research include:

- ◆ New approaches to solving problems, such as the identification of different features in a DNA sequence, the analysis of gene expression and

regulation, the elucidation of protein structure and protein-protein interactions, the determination of the relationship between genotype and phenotype, and the identification of the patterns of genetic variation in populations and the processes that produced those patterns

- ◆ Reusable software modules to facilitate interoperability
- ◆ Methods to elucidate the effects of environmental (non-genetic) factors and of gene-environment interactions on health and disease
- ◆ New ontologies to describe different data types
- ◆ Improved database technologies to facilitate the integration and visualization of different data types, for example, information about pathways, protein structure, gene variation, chemical inhibition and clinical information/phenotypes
- ◆ Improved knowledge management systems and the standardization of data sets to allow the coalescence of knowledge across disciplines

Establishing a true understanding of how organized molecular pathways and networks give rise to normal and pathological cellular and organismal phenotypes will require more than large, experimentally derived data sets. Once again, computational investigation will be essential (Box 3), and there will be a greatly increased need for the collection, storage and display of the data in robust databases. By modelling specific pathways and networks, predicting how they affect phenotype, testing hypotheses derived from these models and refining the models based on new experimental data, it should be possible to understand more completely the difference between a 'bag of molecules' and a functioning biological system.

Grand Challenge I-3 Develop a detailed understanding of the heritable variation in the human genome

Genetics seeks to correlate variation in DNA sequence with phenotypic differences (traits). The greatest advances in human genetics have been made for traits associated with variation in a single gene. But most phenotypes, including common diseases and variable responses to pharmacological agents, have a more complex origin, involving the interplay between multiple genetic factors (genes and their products) and non-genetic factors (environmental influences). Unravelling such complexity will require both a complete description of the genetic variation in the human genome and the development of analytical tools for using that information to understand the genetic basis of disease.

Establishing a catalogue of all common variants in the human population, including single-nucleotide polymorphisms (SNPs), small deletions and insertions, and other structural differences, began in earnest several years ago. Many SNPs have been identified²⁸, and most are publicly available (www.ncbi.nlm.nih.gov/SNP). A public collaboration, the International HapMap Project (www.genome.gov/Pages/Research/HapMap), was formed in 2002 to characterize the patterns of linkage disequilibrium and haplotypes across the human genome and to identify subsets of SNPs that capture most of the information about these patterns of genetic variation to enable large-scale genetic association studies. To reach fruition, such studies need more robust experimental (Box 2) and computational (Box 3) methods that use this new knowledge of human haplotype structure²⁹.

A comprehensive understanding of genetic variation, both in humans and in model organisms, would facilitate studies to establish relationships between genotype and biological function. The study of particular variants and how they affect the functioning of specific proteins and protein pathways will yield important new insights about physiological



processes in normal and disease states. An enhanced ability to incorporate information about genetic variation into human genetic studies would usher in a new era for investigating the genetic bases of human disease and drug response (see Grand Challenge II-1).

Grand Challenge I-4 Understand evolutionary variation across species and the mechanisms underlying it

The genome is a dynamic structure, continually subjected to modification by the forces of evolution. The genomic variation seen in humans represents only a small glimpse through the larger window of evolution, where hundreds of millions of years of trial-and-error efforts have created today's bio-

sphere of animal, plant and microbial species. A complete elucidation of genome function requires a parallel understanding of the sequence differences across species and the fundamental processes that have sculpted their genomes into the modern-day forms.

The study of inter-species sequence comparisons is important for identifying functional elements in the genome (see Grand Challenge I-1). Beyond this illuminating role, determining the sequence differences between species will provide insight into the distinct anatomical, physiological and developmental features of different organisms, will help to define the genetic basis for speciation and will facilitate the characterization of mutational processes. This last point deserves particular attention, because mutation both drives long-term evolutionary change and is the underlying cause of inherited disease. The recent finding that mutation rates vary widely across the mammalian genome¹¹ raises numerous questions about the molecular basis for these evolutionary changes. At present, our understanding of DNA mutation and repair, including the important role of environmental factors, is limited.

Genomics will provide the ability to substantively advance insight into evolutionary variation, which will, in turn, yield new insights into the dynamic nature of genomes in a broader evolutionary framework.

Grand Challenge I-5 Develop policy options that facilitate the widespread use of genome information in both research and clinical settings
Realization of the opportunities provided by genomics depends on effective access to the

BOX 4 Training



Meeting the scientific, medical and social/ethical challenges now facing genomics will require scientists, clinicians and scholars with the skills to understand biological systems and to use that information effectively for the benefit of humankind. Adequate training capacity will be required to address the following needs:

◆ **Computational skills** As biomedical research is becoming increasingly data intensive, computational capability is increasingly becoming a critical skill.

◆ **Interdisciplinary skills** Although a good start has been made, expanded interactions will be required between the sciences (biology, computer science, physics, mathematics, statistics, chemistry and engineering), between the basic and the clinical sciences, and between the life sciences, the social sciences and the humanities. Such interactions will be needed at the individual level (scientists, clinicians and scholars will need to be able to bring relevant issues, concerns and capabilities from different disciplines to bear on

their specific research efforts), at a collaborative level (researchers will need to be able to participate effectively in interdisciplinary research collaborations that bring biology together with many other disciplines) and at the disciplinary level (new disciplines will need to emerge at the interfaces between the traditional disciplines).

◆ **Different perspectives** Individuals from minority or disadvantaged populations are significantly under-represented as both researchers and participants in genomics research. This regrettable circumstance deprives the field of the best and brightest from all backgrounds, narrows the field of questions asked, can lessen sensitivity to cultural concerns in implementing research protocols, and compromises the overall effectiveness of the research. Genomics can learn from successful efforts in training individuals from under-represented populations in other areas of science and health (see, for example, www.genome.gov/Pages/Grants/Policies/ActionPlanGuide).

information (such as data about genes, gene variants, haplotypes, protein structures, small molecules and computational models) by a wide range of potential users, including researchers, commercial enterprises, health-care providers, patients and the public. Researchers themselves need maximum access to the data as soon as possible (see 'Data release', below). Use of the information for the development of therapeutic and other products necessarily entails consideration of the complex issues of intellectual property (for example, patenting and licensing) and commercialization. The intellectual property practices, laws and regulations that affect genomics must adhere to the principle of maximizing public benefit, but must also be consistent with more general and longer-established intellectual property principles. Further, because genome research is global, international treaties, laws, regulations, practices, belief systems and cultures also come into play.

Without commercialization, most diagnostic and therapeutic advances will not reach the clinical setting, where they can benefit patients. Thus, we need to develop policy options for data access and for patenting, licensing and other intellectual property issues to facilitate the dissemination of genomics data.

II Genomics to health

Translating genome-based knowledge into health benefits

The sequencing of the human genome, along with other recent and expected achievements in genomics, provides an unparalleled opportunity to advance our understanding of the role of genetic factors in human health and disease, to allow more precise definition of the non-genetic factors involved, and to apply this insight rapidly to the prevention, diagnosis and



treatment of disease. The report by the US National Research Council that originally envisioned the HGP was explicit in its expectation that the human genome sequence would lead to improvements in human health, and subsequent five-year plans reaffirmed this view^{15–17}. But how this will happen has been less clearly articulated. With the completion of the original goals of the HGP, the time is right to develop and apply large-scale genomic strategies to empower improvements in human health, while anticipating and avoiding potential harm.

Such strategies should enable the research community to achieve the following:

- ◆ Identify genes and pathways with a role in health and disease, and determine how they interact with environmental factors.
- ◆ Develop, evaluate and apply genome-

based diagnostic methods for the prediction of susceptibility to disease, the prediction of drug response, the early detection of illness and the accurate molecular classification of disease.

- ◆ Develop and deploy methods that catalyse the translation of genomic information into therapeutic advances.

Grand Challenge II-1 Develop robust strategies for identifying the genetic contributions to disease and drug response. For common diseases, the interplay of multiple genes and multiple non-genetic factors, not a single allele, usually dictates disease susceptibility and response to treatments. Deciphering the role of genes in human health and disease is a formidable problem for many reasons, including impediments to defining biologically valid phenotypes, challenges in identifying and quantifying environmental exposures, technological obstacles to generating sufficient and useful genotypic information, and the difficulties of studying humans. Yet this problem can be solved. Vigorous development of cross-cutting genomic tools to catalyse advances in understanding the genetics of common disease and in pharmacogenomics is needed. Prominent among these will be a detailed haplotype map of the human genome (see Grand Challenge I-3) that can be used for whole-genome association studies of all diseases in all populations, as well as further advances in sequencing and genotyping technology to make such studies feasible (see 'Quantum leaps', below).

More efficient strategies for detecting rare alleles involved in common disease are also needed, as the hypothesis that alleles that increase risk for common diseases are themselves common³⁰ will probably not be universally true. Computational and experimental methods to detect gene–gene and gene–environment interactions, as well as methods allowing interfacing of a variety of relevant databases, are also required (Box 3). By obtaining unbiased assessments of the relative disease risk that particular gene variants contribute, a large longitudinal population-based cohort study, with collection of extensive clinical information and ongoing follow-up, would be profoundly valuable to the study of all common diseases (Box 1). Already, such projects as the UK Biobank (www.ukbiobank.ac.uk), the Marshfield Clinic's Personalized Medicine Research Project (www.mfldclin.edu/pmrp) and the Estonian Genome Project (www.geenivaramu.ee) seek to provide such resources. But if the multiple population groups in the United States and elsewhere in the world are to benefit fully and fairly from such research (see Grand Challenge II-6), a large population-based cohort study that includes full representation of minority populations is also needed.

BOX 5 Ethical, legal and social implications (ELSI)



Today's genomics research and applications rest on more than a decade of valuable investigation into their ethical, legal and social implications. As the application of genomics to health increases along with its social impact, it becomes ever more important to expand on this work. There is an increasing need for focused ELSI research that directly informs policies and practices. One can envisage a flowering of 'translational ELSI research' that builds on the knowledge gained from prior and forthcoming 'basic ELSI research', which would provide knowledge for direct use by researchers, clinicians, policy-makers and the public. Examples include:

- ◆ The development of models of genomics research that use attention to these ELSI issues

for enhancing the research, rather than viewing such issues as impediments

- ◆ The continued development of appropriate and effective genomics research methods and policies that promote the highest levels of science and of protecting human subjects
- ◆ The establishment of crosscutting tools, analogous to the publicly accessible genomic maps and sequence databases that have accelerated other genomics research (examples of such tools might include searchable databases of genomic legislation and policies from around the world, or studies of ELSI aspects of introducing clinical genetic tests)
- ◆ The evaluation of new genetic and genomic tests and technologies, and effective oversight of their implementation, to ensure that only those with confirmed clinical validity are used for patient care

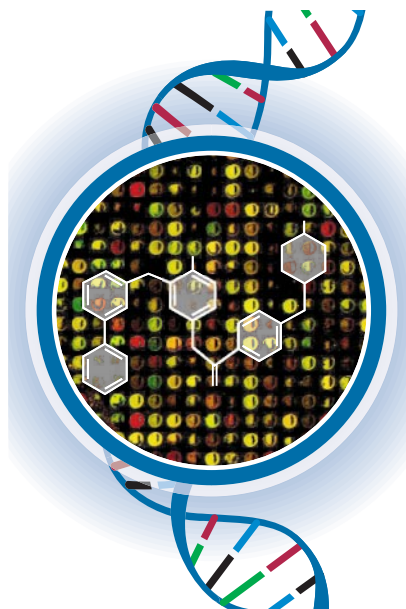
Grand Challenge II-2 Develop strategies to identify gene variants that contribute to good health and resistance to disease

Most human genetic research has traditionally focused on identifying genes that predispose to illness. A relatively unexplored, but important, area of research focuses on the role of genetic factors in maintaining good health. Genomics will facilitate further understanding of this aspect of human biology and allow the identification of gene variants that are important for the maintenance of health, particularly in the presence of known environmental risk factors. One useful research resource would be a 'healthy cohort', a large epidemiologically robust group of individuals (Box 1) with unusually good health, who could be compared with cohorts of individuals with diseases and who could also be intensively studied to reveal alleles protective for conditions such as diabetes, cancer, heart disease and Alzheimer's disease. Another promising approach would be rigorous examination of genetic variants in individuals at high risk for specific diseases who do not develop them, such as sedentary, obese smokers without heart disease or individuals with *HNPCC* mutations who do not develop colon cancer.

Grand Challenge II-3 Develop genome-based approaches to prediction of disease susceptibility and drug response, early detection of illness, and molecular taxonomy of disease states

The discovery of variants that affect risk for disease could potentially be used in individualized preventive medicine — including diet, exercise, lifestyle and pharmaceutical intervention — to maximize the likelihood of staying well. For example, the discovery of variants that correlate with successful outcomes of drug therapy, or with unfortunate side effects, could potentially be rapidly translated into clinical practice. Turning this vision into reality will require the following: (1) unbiased determination of the risk associated with a particular gene variant, often overestimated in initial studies³¹; (2) technological advances to reduce the cost of genotyping (Box 2; see 'Quantum leaps', below); (3) research on whether this kind of personalized genomic information will actually alter health behaviours (see Grand Challenge II-5); (4) oversight of the implementation of genetic tests to ensure that only those with demonstrated clinical validity are applied outside of the research setting (Box 5); and (5) education of healthcare professionals and the public to be well-informed participants in this new form of preventive medicine (Box 6).

The time is right for a focused effort to understand, and potentially to reclassify, all human illnesses on the basis of detailed molecular characterization. Systematic analyses of somatic mutations, epigenetic modifica-



tions, gene expression, protein expression and protein modification should allow the definition of a new molecular taxonomy of illness, which would replace our present, largely empirical, classification schemes and advance both disease prevention and treatment. The reclassification of neuromuscular diseases³² and certain types of cancer³³ provides striking initial examples, but many more such applications are possible.

Such a molecular taxonomy would be the basis for the development of better methods for the early detection of disease, which often allows more effective and less costly treatments. Genomics and other large-scale approaches to biology offer the potential for

developing new tools to detect many diseases earlier than is currently feasible. Such 'sentinel' methods might include analysis of gene expression in circulating leukocytes, proteomic analysis of body fluids, and advanced molecular analysis of tissue biopsies. An example would be the analysis of gene expression in peripheral blood leukocytes to predict drug response. A focused effort to use a genomic approach to characterize serum proteins exhaustively in health and disease might also be highly rewarding.

Grand Challenge II-4 Use new understanding of genes and pathways to develop powerful new therapeutic approaches to disease

Pharmaceuticals on the market target fewer than 500 human gene products³⁴. Even though not all of the 30,000 or so human protein-coding genes⁷ will have products targetable for drug development, this suggests that there is an enormous untapped pool of human gene-based targets for therapeutic intervention. In addition, the new understanding of biological pathways provided by genomics (see Grand Challenge I-2) should contribute even more fundamentally to therapeutic design.

The information needed to determine the therapeutic potential of a gene generally overlaps heavily with the information that reveals its function. The success of imatinib mesylate (Gleevec), an inhibitor of the BCR-ABL tyrosine kinase, in treating chronic myelogenous leukaemia relied on a detailed molecular understanding of the disease's genetic cause³⁵. This example offers promise that therapies based on genomic informa-

BOX 6 Education



Marked health improvements from integrating genomics into individual and public health care depend on the effective education of health professionals and the public about the interplay of genetic and environmental factors in health and disease. Health professionals must be knowledgeable about genomics to use the outcomes of genomics research effectively. The public must be knowledgeable to make informed decisions about participation in genomics research and to incorporate the findings of such research into their own health care. Both groups must be knowledgeable to engage profitably in discussion and decision-making about the societal implications of genomics.

Promising models for genomics and genetics education exist (see, for example, www.nchpeg.org), but they must be expanded and new models developed. We have entered a unique 'educable era' regarding genomics; health professionals and the public are increasingly interested in learning about genomics, but its widespread application to health is still several

years away. For genomics-based health care to be maximally effective once it is widely feasible, and for members of society to make the best decisions about the uses of genomics, we must take advantage now of this unique opportunity to increase understanding. Some examples are:

- ◆ Health professionals vary, both individually and by discipline, in the amount and type of genomics education that they require. So multiple models of effective genomics-related education are needed.

- ◆ Print, web and video educational products that the public can consume when actively seeking genomic information should be created and made easily available.

- ◆ The media are crucial sources of information about genomics and its societal implications. Initiatives to provide the media with greater understanding of genomics are needed.

- ◆ High-school students will be both the users of genomic information and the genomics researchers of the future. Especially as they educate all sectors of society, high-school educators need information and materials about genomics and its implications for society, to use in their classrooms.

tion will be particularly effective. Grand Challenge I-1 describes the 'functionation' of the genome, which will increasingly be the critical first step in the development of new therapeutics. But stimulating basic scientists to approach biomedical problems with a genomic attitude is not enough. A therapeutic mindset, lacking in much of academic biomedical research and training, must be explicitly encouraged, and tools developed and provided for its implementation.

A particularly promising example of the gene-based approach to therapeutics is the application of 'chemical genomics'²⁵. This strategy uses libraries of small molecules (natural compounds, aptamers or the products of combinatorial chemistry) and high-throughput screening to advance understanding of biological pathways and to identify compounds that act as positive or negative regulators of individual gene products, pathways or cellular phenotypes. Although the pharmaceutical industry applies this approach widely as the first step in drug development, few academic investigators have access to this methodology or are familiar with its use.

Providing such access more broadly, through one or more centralized facilities, could lead to the discovery of a host of useful probes for biological pathways that would serve as new reagents for basic research and/or starting points for the development of new therapeutic agents (the 'hits' from such library screens will generally require medicinal chemistry modifications to yield therapeutically usable compounds).

Also needed are new, more powerful technologies for generating deep molecular libraries, especially ones tagged to allow the ready determination of precise molecular targets. A centralized database of screening results should lead to further important biological insights. Generating molecular probes for exploring the basic biology of health and disease in academic laboratories would not supplant the major role of biopharmaceutical companies in drug development, but could contribute to the start of the drug development pipeline. The private sector would doubtless find many of these molecular probes of interest for further exploration through optimization by medicinal chemistry, target validation, lead compound identification, toxicological studies and, ultimately, clinical trials.

Academic pursuit of this first step in drug development could be particularly valuable for the many rare mendelian diseases, in which often the gene defect is known but the small market size limits the private sector's motivation to shoulder the expense of effective pharmaceutical development. Such translational research in academic laboratories, combined with incentives such as the US Orphan Drug Act, could profoundly increase the availability of effective treat-



ments for rare genetic diseases in the next decade. Further, the development of therapeutic approaches to single-gene disorders might provide valuable insights into applying genomics to reveal the biology of more common disorders and developing more effective treatments for them (in the way that, for example, the search for compounds that target the presenilins has led to general therapeutic strategies for late-onset Alzheimer's disease³⁶).

Grand Challenge II-5 Investigate how genetic risk information is conveyed in clinical settings, how that information influences health strategies and behaviours, and how these affect health outcomes and costs

Understanding how genetic factors affect health is often cited as a major goal of genomics, on the assumption that applying such understanding in the clinical setting will improve health. But this assumption actually rests on relatively few examples and data, and more research is needed to provide sufficient guidance about how to use genomic information optimally for improving individual or public health.

Theoretically, the steps by which genetic risk information would lead to improved health are: (1) an individual obtains genome-based information about his/her own health risks; (2) the individual uses this information to develop an individualized

prevention or treatment plan; (3) the individual implements that plan; (4) this leads to improved health; and (5) healthcare costs are reduced. Scrutiny of these assumptions is needed, both to test them and to determine how each step could best be accomplished in different clinical settings.

Research is also required that critically evaluates new genetic tests and interventions in terms of parameters such as benefits, access and cost. Such research should be interdisciplinary and use the tools and expertise of many fields, including genomics, health education, health behaviour research, health outcomes research, healthcare delivery analysis, and healthcare economics. Some of these fields have historically paid little attention to genomics, but high-quality research of this sort could provide important guidance in clinical decision-making — as the work of several disciplines has already been helpful in caring for people with an increased risk of colon cancer as a result of mutations in *FAP* or *HNPCC*³⁷.

Grand Challenge II-6 Develop genome-based tools that improve the health of all

Disparities in health status constitute a significant global issue, but can genome-based approaches to health and disease help to reduce this problem? Social and other environmental factors are major contributors to health disparities; indeed, some would question whether heritable factors have any significant role. But population differences in allele frequencies for some disease-associated variants could be a contributing factor to certain disparities in health status, so incorporating this information into preventive and/or public-health strategies would be beneficial. Research is needed to understand the relationship between genomics and health disparities by rigorously evaluating the diverse contributions of socioeconomic status, culture, discrimination, health behaviours, diet, environmental exposures and genetics.

It is also important to explore applications of genomics in the improvement of health in the developing world (www3.who.int/whosis/genomics/genomics_report.cfm), where both human and non-human genomics will play significant roles. If we take malaria as an example, a better understanding of human genetic factors that influence susceptibility and response to the disease, and to the drugs used to treat it, could have a significant global impact. So too could a better understanding of the malarial parasite itself and of its mosquito vector, which the recently reported genome sequences^{38,39} should provide. It will be necessary to determine the appropriate roles of governmental and non-governmental organizations, academic institutions, industry and individuals to ensure that genomics produces clinical benefits for resource-poor nations, and is used to produce robust local research expertise.

The non-coding part of the human genome is functionally important, yet little is known about it.

To ensure that genomics research benefits all, it will be critical to examine how genomics-based health care is accessed and used. What are the barriers to equitable access, and how can they be removed? This is relevant not only in resource-poor nations, but also in wealthier countries where segments of society, such as indigenous populations, the uninsured, or rural and inner city communities, have traditionally not received adequate health care.

III Genomics to society

Promoting the use of genomics to maximize benefits and minimize harms

Genomics has been at the forefront of giving serious attention, through scholarly research and policy discussions, to the impact of science and technology on society. Although the major benefits to be realized from genomics are in the area of health, as described above, genomics can also contribute to other aspects of society. Just as the HGP and related developments have spawned new areas of research in basic biology and in health, they have also created opportunities for research on social issues, even to the extent of understanding more fully how we define ourselves and each other.

In the next few years, society must not only continue to grapple with numerous questions raised by genomics, but must also formulate and implement policies to address many of them. Unless research provides reliable data and rigorous approaches on which to base such decisions, those policies will be ill-informed and could potentially compromise us all. To be successful, this research must encompass both 'basic' investigations that develop conceptual tools and shared vocabularies, and more 'applied', 'translational' projects that use these tools to explore and define appropriate public-policy options that incorporate diverse points of view.

As it has in the past, such research will continue to have important ramifications for all three major themes of the vision presented here. We now address research that focuses on society itself, more than on biology or health. Such efforts should enable the research community to:

- ◆ Analyse the impact of genomics on concepts of race, ethnicity, kinship, individual and group identity, health, disease and 'normality' for traits and behaviours.
- ◆ Define policy options, and their potential consequences, for the use of genomic information and for the ethical boundaries around genomics research.

Grand Challenge III-1 Develop policy options for the uses of genomics in medical and non-medical settings

Surveys have repeatedly shown that the public is highly interested in the concept that personal genetic information might



guide them to better health, but is deeply concerned about potential misuses of that information (see www.publicagenda.org/issues/pcc_detail.cfm?issue_type=medical_research&list=7). Topping the list of concerns is the potential for discrimination in health insurance and employment. A significant amount of research on this issue has been done⁴⁰, policy options have been published^{41–43}, and many US states have now passed anti-discrimination legislation (see www.genome.gov/Pages/PolicyEthics/Leg/StateIns and www.genome.gov/Pages/PolicyEthics/Leg/StateEmploy). The US Equal Employment Opportunity Commission has ruled that the Americans with Disabilities Act should apply to discrimination based on predictive genetic information⁴⁴, but the legal status of that construct remains in some doubt. Although an executive order protects US government employees against genetic discrimination, this does not apply to other workers. Thus, many observers have concluded that effective federal legislation is needed, and the US Congress is currently considering such a law.

Making certain that genetic tests offered to the public have established clinical validity and usefulness must be a priority for future research and policy making. In the United States, the Secretary's Advisory Committee on Genetic Testing extensively reviewed this area and concluded that further oversight is needed, asking the Food and Drug Administration

to review new predictive genetic tests prior to marketing (www4.od.nih.gov/oba/sacgt/reports/oversight_report.pdf). That recommendation has not yet been acted on; meanwhile, numerous websites offering unvalidated genetic tests directly to the public, often combined with the sale of 'nutraceuticals' and other products of highly questionable value, are proliferating.

Many issues currently swirl around the proper conduct of genetic research involving human subjects, and further work is needed to achieve a satisfactory balance between the protection of research participants from harm and the ability to conduct clinical research that benefits society as a whole. Much effort has gone into developing appropriate guidelines for the use of stored tissue specimens (www.georgetown.edu/research/nrcbl/nbac/hbm.pdf), for community consultation when conducting genetic research with identifiable populations (www.nigms.nih.gov/news/reports/community_consultation.html#exec), and for the consent of non-examined family members when conducting pedigree research (www.nih.gov/sigs/bioethics/nih_third_party_rec.html), but confusion still remains for many investigators and institutional review boards.

The use of genomic information is not limited to the arenas of biology and of health, and further research and development of policy options is also needed for the many other applications of such information. The array of additional users is likely to include the life, disability and long-term care insurance industries, the legal system, the military, educational institutions and adoption agencies. Although some of the research informing the medical uses of genomics will be useful in broader settings, dedicated research outside the healthcare sphere is needed to explore the public values that apply to uses of genomics other than for health care and their relationship to specific contextual applications. For example, should genetic information on predisposition to hyperactivity be available in the future to school officials? Or should genetic information about behavioural traits be admissible in criminal or civil proceedings? Genomics also provides greater opportunity to understand ancestral origins of populations and individuals, which raises issues such as whether genetic information should be used for defining membership in a minority group.

Because uses of genomics outside the healthcare setting will involve a significantly broader community of stakeholders, both research and policy development in this area must involve individuals and organizations besides those involved in the medical applications of genomics. But many of the same perspectives essential to research and policy development for the medical uses of genomics are also essential. Both the potential users of non-medical applications of genomics and the

It should be possible to understand the difference between a 'bag of molecules' and a biological system.

public need education to understand better the nature and limits of genomic information (Box 6) and to grasp the ethical, legal and social implications of its uses outside health care (Box 5).

Grand Challenge III-2 Understand the relationships between genomics, race and ethnicity, and the consequences of uncovering these relationships

Race is a largely non-biological concept confounded by misunderstanding and a long history of prejudice. The relationship of genomics to the concepts of race and ethnicity has to be considered within complex historical and social contexts.

Most variation in the genome is shared between all populations, but certain alleles are more frequent in some populations than in others, largely as a result of history and geography. Use of genetic data to define racial groups, or of racial categories to classify biological traits, is prone to misinterpretation. To minimize such misinterpretation, the biological and sociocultural factors that interrelate genetics with constructs of race and ethnicity need to be better understood and communicated within the next few years.

This will require research on how different individuals and cultures conceive of race, ethnicity, group identity and self-identity, and what role they believe genes or other biological factors have. It will also require a critical examination of how the scientific community understands and uses these concepts in designing research and presenting findings, and of how the media report these. Also necessary is widespread education about the biological meaning and limitations of research findings in this area (Box 6), and the formulation and adoption of public-policy options that protect against genomics-based discrimination or maltreatment (see Grand Challenge III-1).

Grand Challenge III-3 Understand the consequences of uncovering the genomic contributions to human traits and behaviours

Genes influence not only health and disease, but also human traits and behaviours. Science is only beginning to unravel the complicated pathways that underlie such attributes as handedness, cognition, diurnal rhythms and various behavioural characteristics. Too often, research in behavioural genetics, such as that regarding sexual orientation or intelligence, has been poorly designed and its findings have been communicated in a way that oversimplifies and overstates the role of genetic factors. This has caused serious problems for those who have been stigmatized by the suggestion that alleles associated with what some people perceive as 'negative' physiological or behavioural traits are more frequent in certain populations. Given this history and the real potential for recurrence,



it is particularly important to gather sufficient scientifically valid information about genetic and environmental factors to provide a sound understanding of the contributions and interactions between genes and environment in these complex phenotypes.

It is also important that there be robust research to investigate the implications, for both individuals and society, of uncovering any genomic contributions that there may be to traits and behaviours. The field of genomics has a responsibility to consider the social implications of research into the genetic contributions to traits and behaviours, perhaps an even greater responsibility than in other areas where there is less of a history of misunderstanding and stigmatization. Decisions about research in this area are often best made with input from a diverse group of individuals and organizations.

Grand Challenge III-4 Assess how to define the ethical boundaries for uses of genomics

Genetics and genomics can contribute understanding to many areas of biology, health and life. Some of these human applications are controversial, with some members of the public questioning the propriety of their scientific exploration. Although freedom of scientific inquiry has been a cardinal feature of human progress, it is not unbounded. It is important for society to define the appropriate and inappropriate uses of genomics. Conversations

The time is right to develop and apply large-scale genomic strategies to improve human health.

between diverse parties based on an accurate and detailed understanding of the relevant science and ethical, legal and social factors will promote the formulation and implementation of effective policies. For instance, in reproductive genetic testing, it is crucial to include perspectives from the disability community. Research should explore how different individuals, cultures and religious traditions view the ethical boundaries for the uses of genomics — for instance, which sets of values determine attitudes towards the appropriateness of applying genomics to such areas as reproductive genetic testing, 'genetic enhancement' and germline gene transfer.

Implementation: the NHGRI's role

The vision for the future of genomics presented here is broad and deep, and its realization will require the efforts of many. Continuation of the extensive collaboration between scientists and between funding sources that characterized the HGP will be essential. Although the NHGRI intends to participate in all the research areas discussed here, it will need to focus its efforts to use its finite resources as effectively as possible. Thus, it will take a major role in some areas, actively collaborate in others, and have only a supporting role in yet others. The NHGRI's priorities and areas of emphasis will also evolve as milestones are met and new opportunities arise.

The approach that has characterized genomics and led to the success of the HGP — an initial focus on technology development and feasibility studies, followed by pilot efforts to learn how to apply new strategies and technologies efficiently on a larger scale, and then implementation of full-scale production efforts — will continue to be at the heart of the NHGRI's priority-setting process. The following are areas of high interest, not listed in priority order.

Large-scale production of genomic data sets

The NHGRI will continue to support genomic sequencing, focusing on the genomes of mammals, vertebrates, chordates and invertebrates; other funders will support the determination of additional genome sequences from microbes and plants. With current technology, the NHGRI could support the determination of as much as 45–60 gigabases of genomic DNA sequence, or the equivalent of 15–20 human genomes, over the next five years. But as the cost of sequencing continues to decrease, the cost/benefit ratio of sequence generation will improve, so that the actual amount of sequencing done will be greatly affected by the development of improved sequencing technology.

The decisions about which genomes to sequence next will be based on the results of comparative analyses that reveal the ability of genomic sequences from unexplored phylogenetic positions to inform the interpreta-

tion of the human sequence and to provide other insights. Finally, the degree to which any new genomic sequence is completed — finished, taken to an advanced draft stage or lightly sampled — will be determined by the use for which the sequence is generated. And, of course, the NHGRI's sequencing programme will maintain close contact with, and take account of the plans and output of, other sequencing programmes, as has happened throughout the HGP.

A second data set ready for production-level effort is the human haplotype map (HapMap). This project, a collaboration between the NHGRI, many other NIH institutes, and four international partners, is scheduled for completion within three years. The outcome of the International HapMap Project will significantly shape the future direction of the NHGRI's research efforts in the area of genetic variation.

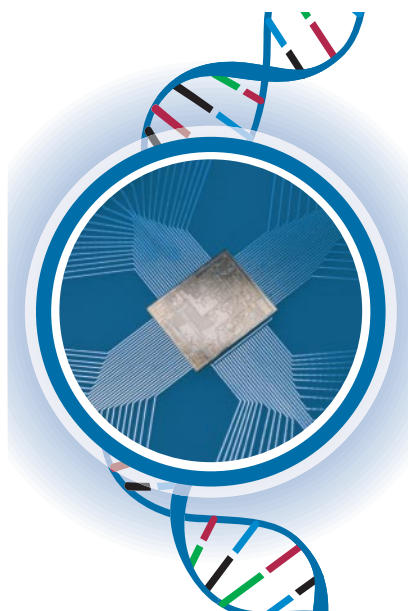
Pilot-scale efforts The NHGRI has initiated the ENCODE Project to begin the development of the human genome 'parts list'. The first phase will address the application and improvement of existing technologies for the large-scale identification of coding sequences, transcription units and other functional elements for which technology is currently available. When the results of the ENCODE Project show evidence of efficacy and affordability at the pilot scale, consideration will be given to implementing the appropriate technologies across the entire human genome.

Technology development Many areas of critical importance to the realization of the genomics-based vision for biomedical research require new technological and methodological developments before pilots and then large-scale approaches can be attempted. Recognizing that technology development is an expensive and high-risk undertaking, the NHGRI is nevertheless committed to supporting and fostering technology development in many of these crucial areas, including the following.

DNA sequencing. There is still great opportunity to reduce the cost and increase the throughput of DNA sequencing, and to make rapid, cheap sequencing available more broadly. Radical reduction of sequencing costs would lead to very different approaches to biomedical research.

Genetic variation. Improved genotyping methods and better mathematical methods are necessary to make effective use of information about the structure of variation in the human genome for identifying the genetic contributions to human diseases and other complex traits.

The genome 'parts list'. Beyond coding sequences and transcriptional units, new computational and experimental approaches are needed to allow the comprehensive deter-



mination of all sequence-encoded functional elements in genomes.

Proteomics. In the short term, the NHGRI expects to focus on the development of appropriate, scalable technologies for the comprehensive analysis of proteins and protein machines in human health and in both rare and complex diseases.

Pathways and networks. As a complement to the development of the genome 'parts list' and increasingly effective approaches to proteome analysis, the NHGRI will encourage the development of new technologies that generate a synthetic view of genetic regulatory networks and interacting protein pathways.

Genetic contributions to health, disease and drug response. The NHGRI will place a high priority on creating and applying new crosscutting genomics tools, technologies and strategies needed to identify the genetic bases of medically relevant phenotypes. Research on the genetic contributions to rare and common diseases, and to drug response, will typically involve biological systems and diseases of primary interest to other NIH institutes and other funding organizations. Accordingly, the NHGRI expects that its involvement in this area of research will often be implemented through partnerships and collaborations. The NHGRI is particularly interested in stimulating research approaches to the identification of gene variants that confer disease resistance and other manifestations of 'good health'.

Society must formulate policies to address many of the questions raised by genomics.

Molecular probes, including small molecules and RNA-mediated interference, for exploring basic biology and disease. Exploration of the feasibility of expanding chemical genomics in the academic and public sectors, particularly with regard to the establishment of one or more centralized facilities, will be pursued by the NHGRI in partnership with others.

Databases Another type of community resource for the biological and biomedical research communities is represented by databases (Box 3). But their support represents a potentially significant problem. Funding agencies, reflecting the interest of the research community, tend to prefer to use their research funds to support the generation of new data, and the ongoing need for continued and increasing support for the data archives and robust access to them is often given less attention. Both the scientific community and the funding agencies must recognize that investment in the creation and maintenance of effective databases is as important a component of research funding as data generation. The NHGRI has been a major source of support for several major genetics/genomics-oriented databases, including the Mouse Genome Database (www.informatics.jax.org/mgihome/MGD/aboutMGD.shtml), the *Saccharomyces* Genome Database (genome-www.stanford.edu/Saccharomyces), FlyBase (flybase.bio.indiana.edu), WormBase (www.wormbase.org) and Online Mendelian Inheritance in Man (www.ncbi.nlm.nih.gov/omim). The NHGRI will continue to be a leader in exploring effective solutions to the issues of integrating, displaying and providing access to genomic information.

Ethical, legal and social research The NHGRI's ELSI research activities will increasingly focus on fundamental, widely relevant, societal issues. The community of scholars and researchers working in these social fields, as well as the scope of issues being explored, need to be expanded. The ELSI research community must include individuals from minority and other communities that may be disproportionately affected by the use or misuse of genetic information. New mechanisms for promoting dialogue and collaboration between the ELSI researchers and genomic and clinical researchers need to be developed; such examples might include structural rewards for interdisciplinary research, intensive summer courses or mini-fellowships for cross-training, and the creation of centres of excellence in ELSI studies to allow sustained interdisciplinary collaboration.

Longitudinal population cohort(s) This promising research resource will be so broadly applicable, and will require such

extensive funding that, although the NHGRI might have a supporting role in design and oversight, success will demand the involvement and support of many other funding sources.

Non-genetic factors in health and disease A consequence of an improved definition of the genetic factors underlying human health and disease will be an improvement in the recognition and definition of the environmental and other non-genetic contributions to those traits. This is another area in which the NHGRI will be involved through the development of new strategies and by forming partnerships.

Use of genomic information to improve health care The NHGRI will catalyse collaboration between the diverse scholarly disciplines whose joint efforts will be necessary for research on the best ways for patients and healthcare providers to make effective use of personalized genetic information in the improvement of health. The NHGRI will also strive to ensure that research in this area is informed by, and extends knowledge of, the societal implications of genomics.

Improving the health of all people It will be important for the NHGRI to support research that explores how to ensure that genomic information is used, to the extent that such information is relevant, to reduce global health disparities. That will include a vigorous effort to increase the representation of minorities in the ranks of genomics researchers. But the full solution of the health disparities problem can only come about through a committed and sustained effort by governments, medical systems and society.

Policy development The NHGRI will continue to help facilitate public-policy development in the area of genetic/genomic science. Effective policy development will require attention to those issues for which it could have the greatest impact on the policy agenda and could help to facilitate genomic science. The NHGRI will also focus on issues that would assist the public in benefiting from genomics, such as privacy of genetic information, access to genetics services, direct-to-consumer/providers marketing, patenting and licensing of genetic information, appropriate treatment of human participants in research, and standards, usefulness and quality in genetic testing.

Data release

An important lesson of the HGP has been the benefit of immediately releasing data from large-scale sequencing projects, as embodied in the Bermuda principles (www.gene.ucl.ac.uk/hugo/bermuda.htm). Some other large-scale data production



projects have followed suit (such as those for full-length cDNAs and single-nucleotide polymorphisms), to the benefit of the scientific community. Scientific progress and public benefit will be maximized by early, open and continuing access to large data sets and by ensuring that excellent scientists are attracted to the task of producing more resources of this sort. For this system to continue to work, the producers of community-resource data sets have an obligation to make the results of their efforts rapidly available for free and unrestricted use by the scientific community, and resource users have an obligation to recognize and respect the important contribution made by the scientists who contribute their time and efforts to resource production.

Although these principles have been generally realized in the case of genomic DNA sequencing, they have not been for many other types of community-resource projects (structural biology coordinates or gene expression data, for example). The development of effective systems for achieving the rapid release of data without restrictions and for providing continued widespread access to materials and research tools should be an integral component of the planning and development of new community resources. The scientific community should also develop incentives to support the voluntary release of such data before publication by individual investigators, by appropriately rewarding and protecting the interests

Scientific progress will be maximized by early, open and continuing access to large data sets.

of scientists who wish to share their data with the community in such a generous manner.

Quantum leaps

It is interesting to speculate about potential revolutionary technical developments that might enhance research and clinical applications in a fashion that would rewrite entire approaches to biomedicine. The advent of the polymerase chain reaction, large-insert cloning systems and methods for low-cost, high-throughput DNA sequencing are examples of such advances that have already occurred.

During the course of the NHGRI's planning discussions, other ideas were raised about analogous 'technological leaps' that seem so far off as to be almost fictional but which, if they could be achieved, would revolutionize biomedical research and clinical practice.

The following is not intended to be an exhaustive list, but to provoke creative dreaming:

- ◆ the ability to determine a genotype at very low cost, allowing an association study in which 2,000 individuals could be screened with about 400,000 genetic markers for \$10,000 or less;
- ◆ the ability to sequence DNA at costs that are lower by four to five orders of magnitude than the current cost, allowing a human genome to be sequenced for \$1,000 or less;
- ◆ the ability to synthesize long DNA molecules at high accuracy for \$0.01 per base, allowing the synthesis of gene-sized pieces of DNA of any sequence for between \$10 and \$10,000;
- ◆ the ability to determine the methylation status of all the DNA in a single cell; and
- ◆ the ability to monitor the state of all proteins in a single cell in a single experiment.

Conclusions

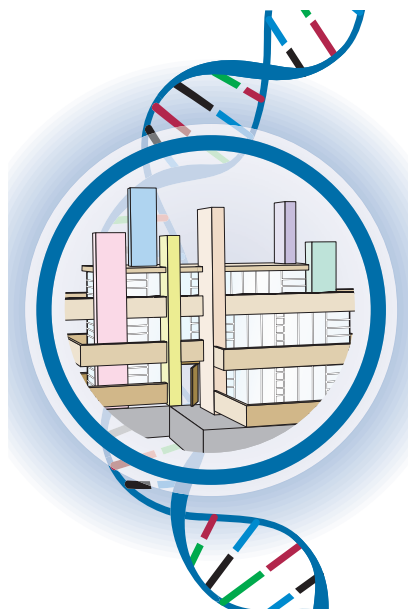
Preparing a vision for the future of genomics research has been both daunting and exhilarating. The willingness of hundreds of experts to volunteer their boldest and best ideas, to step outside their areas of self-interest and to engage in intense debates about opportunities and priorities, has added a richness and audacity to the outcome that was not fully anticipated when the planning process began. To the extent that this article captures the sense of excitement of the new discipline of genomics, it is to their credit. A complete list of the participants in this planning process can be found at www.genome.gov/About/Vision/Acknowledgements.

A final word is appropriate about the breadth of the vision articulated here. A choice had to be made between portraying a broad view of the future of genomics

research and focusing more narrowly on the specific role of the NHGRI. Recognizing that researchers and the public are more interested in the promise of the field than about the funding source responsible, we have focused here on the broad landscape of scientific opportunity. We have, however, identified the areas that are particularly appropriate for leadership by the NHGRI throughout this article. These are generally research areas that are not specific to a particular disease or organ system, but have broader biomedical and/or social implications. Yet even in those instances, the word 'partnership' appears numerous times intentionally. We expect to have partnerships not only with other public funding sources, such as the other 26 NIH institutes and centres, but also with many other governmental agencies, private foundations and private-sector organizations. Indeed, public-private partnerships, such as the SNP Consortium, the Mouse Sequencing Consortium and the International HapMap Project, provide powerful new models for the generation of public data sets with immediate and far-reaching value. Thus, many of the most exciting opportunities in genomics research cross traditional boundaries of specific disease definitions, classically defined scientific disciplines, funding sources and public versus private enterprise. The new era will flourish best in an environment where such traditional boundaries become ever more porous.

Although the opportunities described here are thought to be highly achievable, the formal initiation of specific programmes will require more detailed analysis. The relative priorities of each component must be addressed in the light of limited resources to support research. The NHGRI plans to release a revised programme announcement and other grant solicitations later this year, providing more specific guidance to extramural researchers about plans for the implementation of this vision. Furthermore, in genomics research, we have learned to expect the unexpected. From past experience, it would be surprising (and rather disappointing) if biological, medical and social contexts did not change in unpredictable ways. That reality requires that this vision be revisited on a regular basis.

In conclusion, the successful completion this month of all of the original goals of the HGP emboldens the launch of a new phase for genomics research, to explore the remarkable landscape of opportunity that now opens up before us. Like Shakespeare, we are inclined to say, "what's past is prologue" (*The Tempest*, Act II, Scene 1). If we, like bold architects, can design and build this unprecedented and noble structure, resting on the firm bedrock foundation of the HGP (Figure 2), then the true promise of genomics research for benefiting humankind can be realized.



"Make no little plans; they have no magic to stir men's blood and probably will themselves not be realized. Make big plans; aim high in hope and work, remembering that a noble, logical diagram once recorded will not die, but long after we are gone will be a living thing, asserting itself with ever-growing insistency" (attributed to Daniel Burnham, architect).

Francis S. Collins, Eric D. Green, Alan E. Guttmacher and Mark S. Guyer are at the National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

- Mendel, G. Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines, Abhandlungen*, Brinn 4, 3–47 (1866).
- Avery, O. T., MacLeod, C. M. & McCarty, M. Studies of the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* Type III. *J. Exp. Med.* **79**, 137–158 (1944).
- Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737 (1953).
- Nirenberg, M. W. The genetic code: II. *Sci. Am.* **208**, 80–94 (1963).
- Jackson, D. A., Symons, R. H. & Berg, P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **69**, 2904–2909 (1972).
- Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl Acad. Sci. USA* **70**, 3240–3244 (1973).
- The International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Sanger, F. & Coulson, A. R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**, 441–448 (1975).
- Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA* **74**, 560–564 (1977).
- Smith, L. M. *et al.* Fluorescence detection in automated DNA-sequence analysis. *Nature* **321**, 674–679 (1986).
- The Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
- The Chipping ForeCast II. *Nature Genet.* **32**, 461–552 (2002).
- Guttmacher, A. E. & Collins, F. S. Genomic medicine — A primer. *N. Engl. J. Med.* **347**, 1512–1520 (2002).
- National Research Council. *Mapping and Sequencing the Human Genome* (National Academy Press, Washington DC, 1988).

- US Department of Health and Human Services, US DOE. *Understanding Our Genetic Inheritance. The US Human Genome Project: The First Five Years*. NIH Publication No. 90-1590 (National Institutes of Health, Bethesda, MD, 1990).
- Collins, F. & Galas, D. A new five-year plan for the US Human Genome Project. *Science* **262**, 43–46 (1993).
- Collins, F. S. *et al.* New goals for the US Human Genome Project: 1998–2003. *Science* **282**, 682–689 (1998).
- Hilbert, D. Mathematical problems. *Bull. Am. Math. Soc.* **8**, 437–479 (1902).
- Aparicio, S. *et al.* Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301–1310 (2002).
- Sidow, A. Sequence first. Ask questions later. *Cell* **111**, 13–16 (2002).
- Zhang, M. Q. Computational prediction of eukaryotic protein-coding genes. *Nature Rev. Genet.* **3**, 698–709 (2002).
- Banerjee, N. & Zhang, M. X. Functional genomics as applied to mapping transcription regulatory networks. *Curr. Opin. Microbiol.* **5**, 313–317 (2002).
- Van der Weyden, L., Adams, D. J. & Bradley, A. Tools for targeted manipulation of the mouse genome. *Physiol. Genomics* **11**, 133–164 (2002).
- Hannon, G. J. RNA interference. *Nature* **418**, 244–251 (2002).
- Stockwell, B. R. Chemical genetics: Ligand-based discovery of gene function. *Nature Rev. Genet.* **1**, 116–125 (2000).
- Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
- Tyson, J. J., Chen, K. & Novak, B. Network dynamics and cell physiology. *Nature Rev. Mol. Cell Biol.* **2**, 908–916 (2001).
- Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
- Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
- Reich, D. E. & Lander, E. S. On the allelic spectrum of human disease. *Trends Genet.* **17**, 502–510 (2001).
- Hirschhorn, J. N., Lohmueller, K., Byrne, E. & Hirschhorn, K. A comprehensive review of genetic association studies. *Genet. Med.* **4**, 45–61 (2002).
- Wagner, K. R. Genetic diseases of muscle. *Neurol. Clin.* **20**, 645–678 (2002).
- Golub, T. R. Genomic approaches to the pathogenesis of hematologic malignancy. *Curr. Opin. Hematol.* **8**, 252–261 (2001).
- Drews, J. & Ryser, S. The role of innovation in drug development. *Nature Biotechnol.* **15**, 1318–1319 (1997).
- Druker, B. J. Imatinib alone and in combination for chronic myeloid leukemia. *Semin. Hematol.* **40**, 50–8 (2003).
- Selkoe, D. J. Alzheimer's disease: genes, proteins, and therapy. *Physiol. Rev.* **81**, 741–66 (2001).
- Lynch, H. T. & de la Chapelle, A. Genomic medicine: hereditary colorectal cancer. *N. Engl. J. Med.* **348**, 919–932 (2003).
- Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
- Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002).
- Anderlik, M. R. & Rothstein, M. A. Privacy and confidentiality of genetic information: What rules for the new science? *Annu. Rev. Genom. Hum. Genet.* **2**, 401–433 (2001).
- Hudson, K. L., Rothenberg, K. H., Andrews, L. B., Kahn, M. J. E. & Collins, F. S. Genetic discrimination and health-insurance — An urgent need for reform. *Science* **270**, 391–393 (1995).
- Rothenberg, K. *et al.* Genetic information and the workplace: Legislative approaches and policy challenges. *Science* **275**, 1755–1757 (1997).
- Fuller, B. P. *et al.* Policy forum: Ethics — privacy in genetics research. *Science* **285**, 1359–1361 (1999).
- Miller, P. S. Is there a pink slip in my genes? *J. Health Care Law Policy* **3**, 225–265 (2000).

Acknowledgements

The formulation of this vision could not have happened without the thoughtful and dedicated contributions of a large number of people. The authors were greatly assisted by Kathy Hudson, Elke Jordan, Susan Vasquez, Kris Wetterstrand, Darryl Leja and Robert Nussbaum. A subcommittee of the National Advisory Council for Human Genome Research, including Wylie Burke, William Gelbart, Eric Juengst, Maynard Olson, Robert Tepper and David Valle, provided a critical sounding board for draft versions of this document. We also thank Aravinda Chakravarti, Ellen Wright Clayton, Raynard Kington, Eric Lander, Richard Lifton and Sharon Terry for serving as working-group chairs at the meeting in November 2002 that refined this document. Finally, we thank the hundreds of individuals who participated as workshop planners and/or participants during this 18-month process.

Genetics and the making of *Homo sapiens*

Sean B. Carroll

Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin 53706, USA

Understanding the genetic basis of the physical and behavioural traits that distinguish humans from other primates presents one of the great new challenges in biology. Of the millions of base-pair differences between humans and chimpanzees, which particular changes contributed to the evolution of human features after the separation of the *Pan* and *Homo* lineages 5–7 million years ago? How can we identify the ‘smoking guns’ of human genetic evolution from neutral ticks of the molecular evolutionary clock? The magnitude and rate of morphological evolution in hominids suggests that many independent and incremental developmental changes have occurred that, on the basis of recent findings in model animals, are expected to be polygenic and regulatory in nature. Comparative genomics, population genetics, gene-expression analyses and medical genetics have begun to make complementary inroads into the complex genetic architecture of human evolution.

What is a man,
If his chief good and market of his time
Be but to sleep and feed? a beast, no more.
Sure, he that made us with such large discourse,
Looking before and after, gave us not
That capability and god-like reason
To fust in us unused.

W. Shakespeare, *Hamlet* IV:iv

What makes modern humans different from the great apes and earlier hominids? In what hominids and when in evolution did important physical traits and behaviours appear? Where in our larger brains do human-specific capabilities reside? These have been long-standing questions in palaeoanthropology and comparative anatomy, since the discovery of Neanderthal skulls and the first studies of great apes in the nineteenth century. Now, the mystery of human origins is expanding beyond the description and history of human traits, towards the genetic mechanisms underlying their formation and evolution. With the characterization of the human genome, and that of our chimpanzee cousin on the way, the quest to discover the genetic basis of the physical and behavioural traits that distinguish us from other apes is rapidly gaining momentum.

Genomes diverge as a function of time, and most of the sequence changes that accumulate between any two related species are selectively neutral or nearly neutral in that they do not contribute to functional or phenotypic differences. The great challenge is to elucidate the number, identity and functions of genes, and the specific changes within them, that have shaped the evolution of traits. This has been accomplished for only a few traits in model systems, so it is a difficult task for human features about which we know little, and an enormous prospect to consider the whole arc of human evolution.

In this article, I will examine both the physical and genetic scope of human evolution and the approaches being used to try to understand it. I will first review the current state of our understanding of human evolution from the viewpoint of the fossil record, comparative anatomy and development. These disciplines point to many key traits to be considered, and define the magnitude and nature of evolutionary change in the human lineage. I will preview the picture of human evolution that we might expect to emerge in view of our current knowledge of the genetic architecture of trait evolution in model systems. I will then examine the variety of methods being used on a genome-wide scale and at the level of individual loci to identify genes that may have contributed to the evolution of key traits. I will discuss some of the crucial methodological challenges in distinguishing causative from potentially large

numbers of candidate loci. Finally, I will address some of the disciplines in which future advances are likely to have a central role in furthering our knowledge of the genetic and developmental basis of human evolution.

Hominid evolution

The hominid tree

To approach the origins of human traits at the genetic level, it is essential to have as a framework a history of our lineage and the characters that distinguish it. It is inadequate and misleading to consider just the comparative anatomy and development (or genomes) of extant humans, chimpanzees and other apes, and then to attempt to infer how existing differences might be encoded and realized. Each of these species has an independent lineage that reaches back as far or further than hominins (‘hominins’ refers to humans and our evolutionary ancestors back to the separation of the human and ape lineages; ‘hominids’ to humans and the African apes) (Fig. 1). The evolution of ‘modern’ traits was not a linear, additive process, and ideas about the tempo, pattern and magnitude of change can only be tested through fossil evidence, which is always subject to revision by new finds. The fossil record continues to shape views of three crucial issues in hominid evolution. First, what distinguishes hominins from the apes? Second, what distinguishes modern humans (*Homo sapiens*) from earlier hominins? And third, what was the nature of the last common ancestor of hominins and the *Pan* lineage?

Inferences about the chronological order and magnitude of the evolution of hominin characters depend on a model of the hominin evolutionary tree. This has always been a contentious issue and remains unsettled, in part because of the pace of exciting new fossil finds over the past two decades. An emerging view portrays hominin evolution as a series of adaptive radiations in which many different branches of the hominin lineage were formed, but died out^{1,2} (Fig. 1). One prediction of this model is that various anatomical features would be found in different combinations in hominins through their independent acquisition, modification and loss in different species. For example, the recent, stunning discovery of a 6–7 million year (Myr)-old fossil cranium of *Sahelanthropus tchadensis* that had a chimpanzee-sized brain but hominin-like facial and dental features³ is the sort of morphology that would be consistent with a radiation of ape-like animals from which the stem of the hominin lineage emerged (although the interpretation of this fossil’s affinities is controversial⁴).

What makes a human?

Evolutionary trends in fossil hominins. Ideally, if every possible fossil human and ape species were identified and many fairly

complete specimens were available, one could reconstruct the emergence of human and chimpanzee features through time. But that is not the case for most lineages; in fact, there are no identified archaic chimpanzee fossils. We must make do with a partial and often confusing picture of human trait evolution. From a rather extensive list of qualitative and quantitative features that distinguish humans from other apes^{5,6} (Box 1), our large brain, bipedalism, small canine teeth, language and advanced tool-making capabilities^{7,8} have been the focus of palaeoanthropology. The major physical traits are generally not singular elements, but entail concomitant changes in skeletal features involved in locomotion (for example, in the vertebral column, pelvis and feet, and in limb proportions), grasping (hand morphology and an opposable,

elongated thumb) and chewing of food (the mandible and dentition), as well as life-history traits such as lifespan. It is fortunate that most of the skeletal features lend themselves to detailed quantitative studies of the fossil record.

Some trends in the evolution of body size, brain size and dentition are evident within the hominins. More recent species are characterized by larger body mass, relatively larger brains, longer legs relative to the trunk and small teeth, whereas earlier species had, in general, smaller brains and bodies (Table 1), shorter legs relative to the trunk and large teeth^{7,9,10}. I highlight these traits to focus attention on the magnitude and timescale of character evolution, and on the (increasing) number of generally recognized hominin taxa. Whatever the branching pattern in the hominin tree, sub-

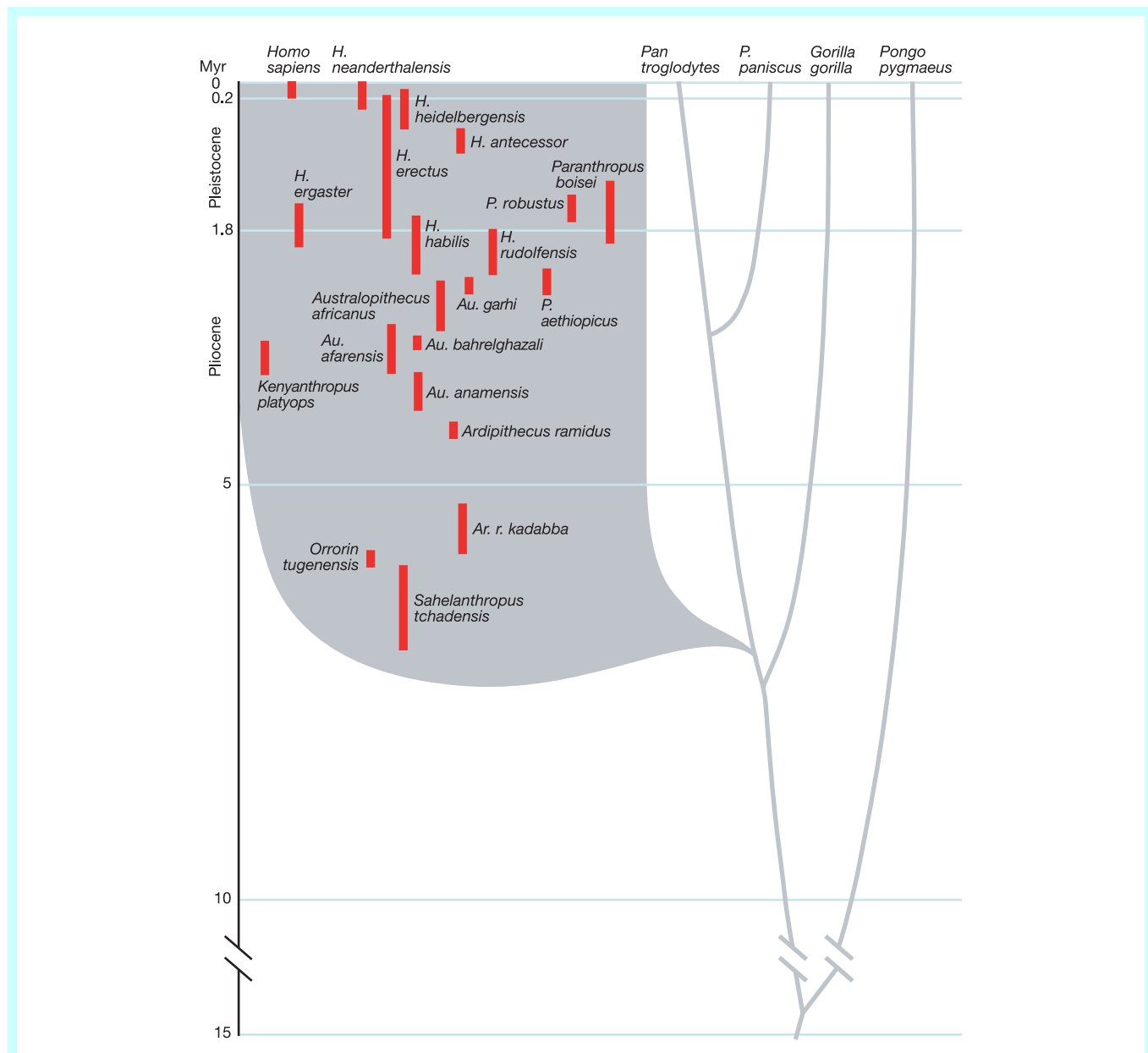


Figure 1 The timescale and phylogeny of hominids. Ape relationships are shown in grey for the chimpanzee (*Pan troglodytes*), bonobo (*P. paniscus*), gorilla and orangutan (*Pongo pygmaeus*). The approximate times of divergence are derived from molecular data (summarized in ref. 89). The phylogenetic relationships among hominins (shaded) are uncertain. The solid red bars denote the time span of the fossil species and/or the uncertainty of fossil ages. The identity of the last common ancestor of chimpanzees and

humans (LCA) is not known. Note that the estimated age of *Sahelanthropus tchadensis* predates molecular estimates of the time of the chimpanzee–human divergence. This species could pre- or postdate the LCA. Also note that *Homo sapiens* represent only the last 3% of the time span of hominin evolution. Hominin distributions and nomenclature are based primarily on refs 1, 90.

stantial relative changes occurred over an extended time span and a significant number of speciation events. There was a marked increase in absolute brain size by the Early Pleistocene and again in the Middle Pleistocene, with a long interval of perhaps 1 Myr during which brain size did not change significantly^{9,11} (Table 1). With regard to modern *H. sapiens*, it is interesting to note that body and brain size were even greater in *H. neanderthalensis*; there is no obvious physical explanation for the success of *H. sapiens* and the demise of *H. neanderthalensis*¹¹.

A beautiful mind: insights from comparative neuroanatomy. The relative increase in brain size, although marked, is only a crude index of a potential increase in cognitive abilities. Because it has long been appreciated that there are discrete areas of the brain that process various cognitive, motor and sensory functions, comparative neuroanatomists have sought to identify areas that might be central to the evolution of human capacities. There is a long-standing notion that the frontal cortex (involved in planning, organization, personality, behaviour and other 'higher' cognitive functions) is disproportionately larger in humans¹², but this now seems not to be the case (it is larger, but not disproportionately so¹³). As gross anatomical differences do not account for cognitive capabilities, relative differences in the size, cellular composition, detailed cytoarchitecture and/or connectivity of human and great ape brain areas have been sought to explain the emergence of human capabilities.

Of paramount interest is the production and understanding of speech. Two areas in particular have commanded the greatest attention. One is Broca's area in the frontal lobe of the neocortex (Fig. 2). This region is larger in the left hemisphere of the brain than in the right, an asymmetry that has been correlated with language ability. From magnetic resonance images of chimpanzees, bonobos and gorillas, a similar left-right asymmetry has been found in these great apes¹⁴ (Fig. 2). This indicates that the neuroanatomical substrate of left-hemisphere dominance in speech production preceded the origin of hominins. The left hemisphere also usually controls right-handedness, so it is interesting to note that in captive apes, manual gestures are right-hand biased, and this bias is increased when vocalization is combined with gesturing¹⁵, indicating a left-hemisphere-controlled communication process.

A second area of interest is Wernicke's posterior receptive language area in the temporal lobe (Fig. 2). A site within this area, the planum temporale, is implicated in human communication (both spoken and gestural) and musical talent, and also shows a left-hemisphere dominance. In most humans, the Sylvian fissure associated with the left planum temporale extends more posteriorly. Evidence for this asymmetry has been found in fossil

endocasts in *H. habilis*, *H. erectus* and *H. neanderthalensis*¹⁶. More importantly, an asymmetrical planum temporale pattern has recently been demonstrated in chimpanzees^{17,18} (Fig. 2).

Several hypotheses have been forwarded to explain the presence of these two human-communication-associated neuroanatomical landmarks in great apes^{14,17}. Although it is possible that they arose and acquired their functions independently in each lineage, the most parsimonious explanation is that the common ancestor of great apes and humans had asymmetrical centres that were involved in communication, and that these structures underwent independent evolutionary modifications in chimpanzees and hominins. If this is the case, then the challenge to comparative neuroanatomy is to identify more subtle differences in suborganization (that is, 'microanatomy') that affect the interconnections of cortical regions, in local circuitry and/or cytoarchitecture^{19,20} that might be unique to human brains. Recently, it has been found that the dimensions of the vertical columns of neurons in the cortex, known as 'mini-columns', differ between humans and chimpanzees in the planum temporale²¹. In addition, area 10 of the prefrontal cortex, which is involved in higher cognitive functions, has been shown to be enlarged and specialized in humans relative to apes²². These observations suggest that human capacities are more a product of quantitative changes in specialized areas than of neuroanatomical novelties.

Development of hominid features

The morphological differences between modern humans, earlier hominins and the great apes are, of course, the product of changes during development. Comparative studies of human and chimpanzee skull development²³, remarkably detailed examinations of Neanderthal craniofacial ontogeny^{22,24–26}, and early hominin dentition formation²⁷ have yielded crucial insights into a variety of developmental shifts that underlie modern human cranial size and morphology.

One of the long-appreciated, fundamental differences in chimpanzee and human development is the relative rate of skull growth and maturation. Human neonates have less mature skulls in terms of shape in comparison to young chimpanzees, but much larger skulls (and brains)^{23,28}. These are classically described as heterochronic changes, which produce neotenic features in which maturation is retarded, size increases and shape resembles juvenile forms of ancestors²⁹. Chimpanzee and human skulls eventually grow to the same size, reflecting further relative shifts in juvenile and adolescent growth periods, and marked differences in face size and cerebral volume. Importantly, all of the skeletal changes associated with bipedalism are structural innovations independent of neoteny. These observations suggest that the human brain is not a product of simple shifts in growth relationships, but of multiple, independent and superimposed modifications.

Box 1

Selected traits that distinguish humans from other apes^{5–7}

- Body shape and thorax
- Cranial properties (brain case and face)
- Relative brain size
- Relative limb length
- Long ontogeny and lifespan
- Small canine teeth
- Skull balanced upright on vertebral column
- Reduced hair cover
- Elongated thumb and shortened fingers
- Dimensions of the pelvis
- Presence of a chin
- S-shaped spine
- Language
- Advanced tool making
- Brain topology

Table 1 Evolution of brain and body size in hominids

Species	Estimated age ^{1,3,7,86–88} (Myr ago)	Body size ⁷ (kg)	Brain size ^{3,7} (cm ³)
<i>Homo sapiens</i>	0–0.2	53	1,355
<i>H. neanderthalensis</i>	0.03–0.25	76	1,512
<i>H. heidelbergensis</i>	0.3–1	62	1,198
<i>H. erectus</i>	0.2–1.9	57	1,016
<i>H. ergaster</i>	1.5–1.9	58	854
<i>H. rudolfensis</i>	1.8–2.4	-	752
<i>H. habilis</i>	1.6–2.3	34	552
<i>Paranthropus boisei</i>	1.2–2.2	44	510
<i>Australopithecus africanus</i>	2.6–3	36	457
<i>Au. afarensis</i>	3–3.6	-	-
<i>Au. anamensis</i>	3.5–4.1	-	-
<i>Ardipithecis ramidus kadabba</i>	5.2–5.8	-	-
<i>Sahelanthropus tchadensis</i>	6–7	-	~320–380

This list does not include all recognized species.

With respect to more recent hominin species, modern human craniofacial form appears to have been shaped by changes in elements that influence the spatial position of the face, neurocranium and cranial base. Modern humans are marked by greater roundedness of the cranial vault and facial retraction (the antero-posterior position of the face relative to the cranial base and neurocranium)²⁵. Comparisons with Neanderthal skull development inferred from fossils of different aged individuals suggest that the characteristic cranial differences between Neanderthal and modern human features arise early in ontogeny^{24,26}.

Our prolonged childhood, delayed sexual maturation and long lifespan are life-history traits that shape important aspects of human society. Insights into the evolution of these development shifts in evolution are offered by detailed comparative analysis of dental development, which is correlated with stages of primate growth and development. Rates of enamel formation in fossil hominins suggest that tooth-formation times were shorter in australopiths and early members of the genus *Homo* than they are in modern humans²⁷. This indicates that the modern pattern of dental formation and correlated developmental traits appeared late in human evolution. When considered in the context of other traits, such as brain size and body proportions, a mosaic pattern of evolution emerges with different traits appearing at different times and perhaps in different combinations in hominin history³⁰.

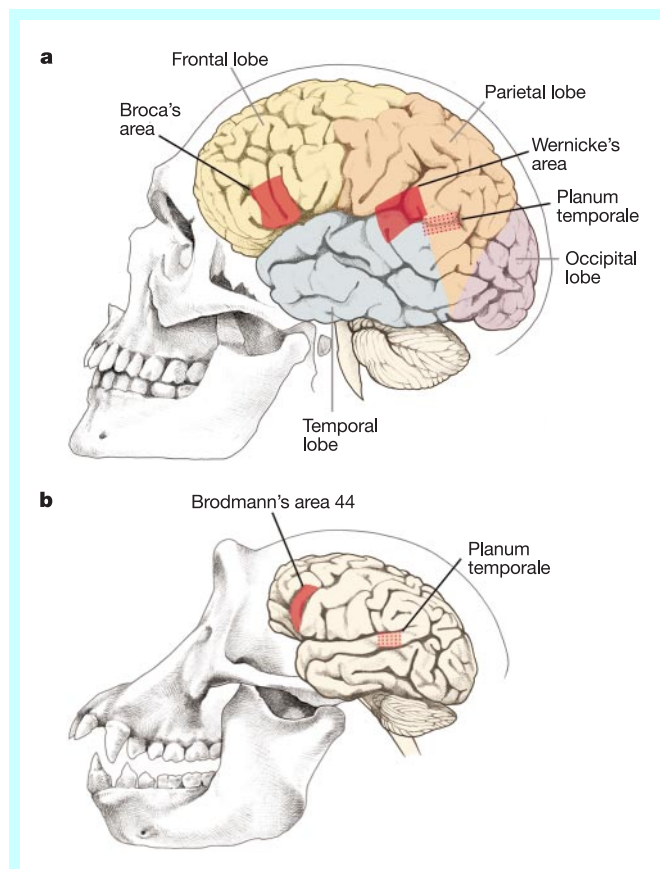


Figure 2 Comparative neuroanatomy of humans and chimpanzees. Lateral views of the left hemispheres of a modern human and a chimpanzee brain. Although the overall skull sizes are roughly comparable, the human cranial capacity and brain are much larger. **a**, Two areas of the human brain that are associated with communication are shown: Broca's area in the frontal lobe and Wernicke's area, which overlaps the posterior temporal lobe and parts of the parietal lobe. In the left hemisphere, Broca's area is larger, as is the planum temporale, which lies below the surface in Wernicke's area. **b**, These asymmetries have been found in corresponding regions of chimpanzee brains^{15,17}, suggesting that the areas in humans might be elaborations of a pre-existing communication centre in a common ancestor of apes and humans.

Was human evolution special?

The magnitude, rate and pattern of change during hominin evolution, inferred from the fossil record, comparative neuroanatomy and embryology, provide the essential foundation for approaching the genetics of human evolution. From the studies discussed above, five key points emerge that have a bearing on attempts to reconstruct the genetic events that underlie the origin and modification of human traits.

First, trait evolution was nonlinear. The $\sim 1,000\text{-cm}^3$ increase in brain size over 5–7 Myr did not occur at the same relative rate in hominin phylogeny: it was static at times, faster in some intervals, and reversed slightly more recently. Second, most trait evolution can be characterized as simple quantitative changes (that is, traits are continuous). Third, evolutionary rates were not at all exceptional with respect to mammalian evolution. For example, fossil horse lineages in the late Pliocene–Pleistocene show similar rates of body-size and other character changes as do those of hominids³¹. Fourth, much evolutionary change preceded the origin of the *Homo* genus and of *H. sapiens*: the history of our species represents just the last 3% of the time span of hominin evolution (Fig. 1). And fifth, many characters are present not only in humans, but also in apes. This suggests that modification of existing structures and developmental pathways, rather than the invention of new features, underlies much of human evolution.

These observations indicate that morphological evolution in hominins was not special, but the product of genetic and developmental changes typical of other mammals and animals.

Genetics of human evolution

Genetic architecture of trait evolution

Given the dimensions of hominin evolution, inferred from the fossil record and comparative anatomy, what can we expect in terms of the genetic complexity underlying trait evolution? For example, there is a long-standing tendency for events that are perceived to be relatively 'rapid' in the fossil record to be ascribed to perhaps one or a few radical mutations³², including recent human evolution^{33,34}. Could the relative increase in brain size over 5 Myr, or its expanded cognitive function, be due to just one or a few genetic changes? The best (and, at present, the only available) guides for this question are detailed genetic studies in model organisms, which have achieved success in dissecting the genetics of complex trait formation, variation and evolution. Six essential general concepts have been established in model systems that pertain to the potential genetic architecture of human trait evolution:

- (1) Variation in continuous, quantitative traits is usually polygenic. Studies of variation in model species^{35–38} reveal that many genes of small effect, and sometimes one or a few genes of large effect, control trait parameters. In humans, a study of variation in 20 anthropometric variables in two different ethnic human populations suggested that more than 50% of variation was polygenic³⁹.
- (2) The rate of trait evolution tells us nothing about the number of genes involved. Studies of artificial selection^{35,38,40} and of inter-specific divergence⁴¹ indicate that the intensity of selection and heritability are more important determinants of evolutionary rate than is the genetic complexity of the traits under selection. There is considerable standing variation in traits, including characters that might be thought of as highly constrained, such as limb morphology in tetrapods⁴². In general, the observed rates of evolution under natural selection are far slower than is potentially possible^{43,44}. Genetic variation or genetic complexity is not the limiting factor³²; indeed, considerable genetic variation underlies even phenotypically invariant traits^{45–48}. Because the rate at which a trait emerges in the fossil record tells us nothing about genetic architecture, the temptation to invoke macromutational models for 'rapid change'^{33,49} must be resisted in the absence of genetic evidence.
- (3) Morphological variation and divergence are associated with

genes that regulate development. Comparisons of the developmental basis of body-pattern evolution in animals suggest that morphological evolution is a product of changes in the spatiotemporal deployment of regulatory genes and the evolution of genetic regulatory networks^{50–52}. Developmental changes in the human lineage are expected to be associated with genes that affect developmental parameters, such as those that encode transcription factors and members of signal-transduction pathways.

(4) Mutations responsible for trait variation are often in non-coding, regulatory regions. When it has been possible to localize variation in genes that underlie phenotypic variation or protein-level differences, insertions or substitutions in regulatory regions and non-coding regions are often responsible^{53–57}.

(5) Multiple nucleotide replacements often differentiate alleles. Fine-scale analysis of quantitative trait loci has often revealed that functional differences between alleles are due to multiple nucleotide differences^{55,56}. It also indicates that non-additive interactions between sites within a locus may be key to the differentiation of alleles, and that the contribution of any individual site may be modest (and difficult to detect).

(6) There is some concordance between genes responsible for intraspecific variation and interspecies divergence. Genetic analyses of interspecies divergence is only possible under certain circumstances, when laboratory breeding can overcome species barriers and traits can be mapped. In some cases, it has been found that some of the same loci are involved in both within-species variation and between-species divergence^{35,58}. This raises some hope that studies of intraspecific variation in humans could lead to genes that have been important in human history.

Since human trait evolution has followed a similar, incremental course as traits studied in model systems, these six concepts suggest that we should expect a highly polygenic basis for complex traits such as brain size, craniofacial morphology and development, cortical speech and language areas, hand and digit morphology, dentition and post-cranial skeletal morphology. We should also anticipate that multiple changes in non-coding regulatory regions and in regulatory genes are of great importance. But how can we find them?

The arithmetic of human sequence evolution

All genetic approaches to human origins are fundamentally comparative, and seek to identify genetic changes that occurred specifically in the human lineage and contributed to the differentiation of humans from our last common ancestor with either apes or other species of *Homo*. Our primary comparative reference is the genome of the chimpanzee (*Pan troglodytes*), our closest living relative, with whom we share a common ancestor that lived 5–7 Myr ago. The arithmetic that sets the problem for human evolutionary genetics is as follows: first, the most extensive comparison of chimpanzee and human genomic sequences indicates an average substitution level of ~1.2% in single-copy DNA⁵⁹; second, the human genome comprises ~3 × 10⁹ base pairs; third, it is reasonable to assume that one-half of the total divergence between chimpanzees and humans occurred in the human lineage (~0.6%); and fourth, this amounts to ~18 × 10⁶ base-pair changes. In addition, there are an unknown number of gene duplications and pseudogene, transposon and repetitive element changes in each lineage. A recent small-scale survey indicated that insertions and deletions (indels) might account for another 3.4% of differences between chimpanzee and human genomes, with the bulk of that figure contributed by larger indels⁶⁰. A good deal of genomic change might be the noise of neutral substitutions and the gain and loss of repetitive elements over long time spans (more than 46% of human DNA is composed of interspersed repeats), but some small fraction of the changes in genomic sequence is responsible for the hereditary differences between species. The crux of the challenge is how to identify specific changes that are biologically meaningful from the many that are not.

In the case of human evolution, there are three basic genetic issues that we would like to grasp. First, how many genes were directly involved in the origin of human anatomy, physiology and behaviour (a few, dozens, hundreds or thousands)? Second, which specific genes contributed to the emergence of particular human traits? And third, what types of change in these genes contributed to evolution (for example, gene duplications, amino-acid replacements or regulatory sequence evolution)? In the few pioneering studies that are directly addressing the genetic basis of human–chimpanzee divergence, different but somewhat complementary strategies are being pursued that are beginning to reveal the scope of human genetic evolution and, in some cases, specific genes that might have been under selection in the course of recent human evolution.

Comparative genomics

The most readily detected differences between animal genomes are expansions or contractions of gene families. Although the full chimpanzee genome is not yet available, a partial comparative map indicates that there are regions of the human genome that might not be represented in chimpanzees or other apes⁶¹. Such regions could be due to duplications or insertions that occurred in the hominin lineage or to deletions in the chimpanzee lineage. One gene family, dubbed *morpheus*, underwent expansion as part of a segmental duplication on human chromosome 16 (ref. 62). This expansion is shared by other great apes, but it seems that there were human lineage-specific duplications as well.

On the basis of comparisons with other genomes, particularly the recently reported draft mouse sequence⁶³, such lineage-specific duplications are expected. In the 75 Myr or more since the divergence of the common ancestor of mice and humans, several dozen clusters of mouse-specific genes arose that are generally represented by a single gene in the human genome⁶³. The shorter divergence time between humans and apes suggests that the human-specific gene set will be smaller. It is interesting to note that a significant fraction of the mouse gene clusters encode proteins with roles in reproduction, immunity and olfaction. This indicates that sexual selection, pathogens and ecology can shape the main differences in coding content between mammals. It should also be noted that 80% of mouse genes have a 1:1 orthologue in the human genome, and that more than 99% have some homologue⁶³. These figures and synteny data suggest that there is a gene repertoire that is qualitatively nearly identical among mammals. The presence or absence of particular gene duplicates might reflect adaptively driven change, but further evidence will be necessary to determine whether positive selection has acted on genes.

Thousands of adaptive changes in the human proteome?

The first place that adaptive genetic changes have been looked for is in the coding sequences for proteins. If the 18 × 10⁶ substitutions in the human lineage are evenly distributed throughout the genome, only a small fraction will be expected to fall within coding regions. Assuming that the average protein is ~400 amino acids in length, and that there are ~30,000 protein-coding genes, only ~3.5 × 10⁷ base pairs (or a little more than 1.5% of the genome) consists of coding regions^{64,65}. So, assuming neutrality and ignoring the selective removal of deleterious changes in protein sequences, ~1.5% of these 18 × 10⁶ substitutions (or 270,000 sites) may contribute to protein evolution. A fraction of these (roughly one-quarter) are synonymous substitutions, so the total number of amino-acid replacements in the human lineage could be of the order of ~200,000. This figure is in good agreement with observed average rates of amino-acid replacement in mammals⁶⁶.

Various methods have been developed to detect whether amino-acid replacements could be the result of positive selection—that is, adaptive evolution^{67,68}. To estimate the extent of positive selection in human protein evolution, Fay *et al.*⁶⁹ surveyed sequence-divergence data for 182 human and Old World monkey genes, and polymorph-

ism data for a similar number of human genes. Taking into consideration the frequency of common polymorphisms (ignoring rare alleles), a greater-than-expected degree of amino-acid replacements was observed, which is evidence of selection. When extrapolated to the entire proteome, 35% of amino-acid substitutions between human and Old World monkeys were estimated to have been driven by positive selection. Applied to human–chimpanzee divergence, this would extrapolate to ~70,000 adaptive substitutions in the human lineage. This figure is substantially larger than would be expected if most mutations were neutral or nearly neutral⁷⁰. If it is even the correct order of magnitude, it forecasts a nightmare for the identification of key genes under selection, because this figure suggests that, on average, two or more adaptive substitutions have occurred in every human protein in the last 5 Myr.

It is possible that the figure, based on the study of less than 0.5% of the human proteome, is an overestimate of the fraction or distribution of adaptive replacements. It is clear that some proteins

are under strong pressure to remain constant, whereas others, especially those involved in so-called ‘molecular arms races’, are under pressure to change. For example, major histocompatibility complex proteins, which interact with diverse and changing foreign substances, show clear signatures of selection⁷¹. Proteins involved in reproduction that play a part in sperm competition or gamete recognition also appear to evolve faster and under some degree of positive selection⁷². A host of human male reproductive proteins have greater-than-average ratios of amino-acid replacements⁷³. Although accelerated protein evolution can also be the consequence of relaxed constraints, the correlation of higher levels of amino-acid replacements in proteins that have a role in reproduction and immunity seems to be biologically and selectively driven.

The population genetics- and protein-sequence-based statistical estimates of adaptive evolution require three caveats regarding how much they tell us about human evolution. First, there are generally no direct functional data that either test or demonstrate whether a human protein is indeed functionally diverged from an ape orthologue. Second, the proteins for which signatures of selection have been detected generally do not affect development. And third, the proteome is just part of the whole picture of genome evolution. Non-coding sequences, including transcriptional *cis*-regulatory elements, the untranslated regions of messenger RNAs, and RNA-splicing signals, contribute considerably to evolution by affecting the time, place and level of gene expression (see above). Ever since the pioneering comparative analysis of ape and human protein-sequence divergence nearly three decades ago⁷⁴, it has generally been anticipated that changes in gene regulation are a more important force than coding-sequence evolution in the morphological and behavioural evolution of hominins.

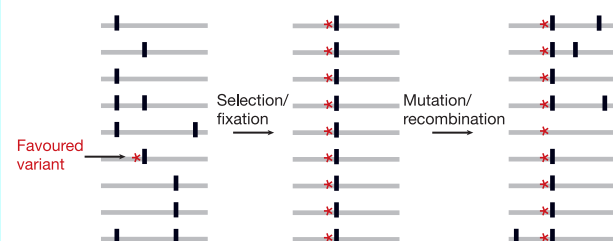
Evolution of human gene expression

How large is the functional compartment of non-coding sequences—the other 98% of the genome? A recent estimate suggests that perhaps twice as much non-coding DNA is under selection than coding DNA⁶³. So, we would also expect a large number of substitutions in the human lineage, of the order of several hundred thousand, with potential functional consequences in non-coding DNA. Even if one applies a much smaller, more conservative estimate of the fraction of adaptive substitutions in non-coding DNA, such as 2% (ref. 75), one still reaches a figure of more than 10,000 adaptive substitutions in human genes and their regulatory regions. The problem is that regulatory sequences are more difficult to analyse: we have no algorithms that can infer biological function from tracts of intergenic or intronic sequence, let alone to decipher how base-pair changes affect function. It is therefore perhaps understandable why non-coding regions have received little attention at the level of population genetics. However, a growing body of work in quantitative genetics and on the evolution of development has shown that regulatory sequences are central to changes in gene expression and morphology. New methodologies have been required to detect the evolution of gene expression and regulatory sequences.

A first step towards the identification of human-specific gene-expression patterns was recently taken by Enard *et al.*⁷⁶, who used genome microarrays to analyse within- and between-species differences in primate gene expression. Analysis of RNA expression profiles from the left prefrontal lobe (Brodmann area 9, which is thought to be involved in cognitive functions) of adult male humans, chimpanzees and an orangutan, and from the neocortex of humans, chimpanzees and macaques, indicated an apparent acceleration of gene-expression differences in the human brain relative to other primates and to other tissues. Protein-expression analyses were also consistent with the idea that relative changes in protein-expression levels were accelerated in the evolution of the human brain, and could be detected for ~30% of the proteins surveyed. The concordance between RNA and protein-level data

Box 2 Selective sweeps

If a change in a gene is favoured, then selection may drive the allele bearing that change to fixation (left and centre of the figure). In the process, neutral variation at linked sites ‘hitchhikes’ along with the selected site; this is known as a ‘selective sweep’. The physical limits of the sweep depend on the strength of the linkage between selected and adjacent sites. After a sweep, variation may again begin to build up, and initially there will be a relatively high frequency of rare polymorphisms (right of the figure). Tajima⁹¹ proposed a statistic (*D*) that tests for selective neutrality. If the frequencies of polymorphisms are skewed, with an excess of rare types, this gives a negative value (neutral value = 0) and can be indicative of a recent selective sweep. Tajima’s *D* is sensitive to other factors apart from selection that can also yield a negative value. A recent expansion in population size from a relatively small population will produce similar patterns of genetic variation and *D* values. In human populations, population history (for example, drift and expansion) and population structure (ethnicity, migration and immigration) will affect *D* values at all loci, whereas selection will affect *D* values at selected and linked loci. The mean *D* values for 437 loci range from –0.69 to –1.25, depending on sampling methods⁹², indicating that population structure and history has had an effect. These negative values underscore a challenge in human evolutionary genetics to distinguish selective sweeps at loci from population-based effects. The *D* value obtained for the human *FOXP2* locus was –2.20, the second largest negative value among all human genes surveyed so far⁸⁴. Other methods have been developed to detect positive selection: for example, by identifying areas of extended haplotype homozygosity⁸¹. It is important to emphasize that all of these methodologies detect signs of recent selection in the *Homo sapiens* lineage. The preceding 5–6 Myr of hominin genetic history, a period when we know from the fossil record that many human features arose, is not addressed by these methods.



indicates that regulatory changes have occurred in a substantial fraction of genes. Indeed, a recent survey of humans heterozygous at 13 loci revealed allelic variation in gene-expression levels at 6 loci⁷⁷. Both intraspecific variations and interspecific divergence in gene expression are probably due to substitutions in non-coding regions that influence transcript or protein abundance through transcriptional or post-transcriptional mechanisms. These data further suggest that quantitative changes in gene expression should be expected as a general feature that accompanies species divergence, and that the raw material for evolutionary changes in gene expression appears to be widely available in non-coding DNA.

The microarray experiments raise many challenges for future progress. Specifically, how can changes that contribute to human anatomy, physiology or behaviour be sorted out from those that don't? Gene-expression data are correlative, not definitive in terms of identifying cause and effect. Many developmental and genetic mechanisms could contribute to the overall pattern observed. For example, a change in the composition of a tissue (for example, the relative proportions of cell types) will be accompanied by altered expression profiles, but many of these changes will be an indirect consequence of a developmental change, not the cause. Similarly, changes in levels or activities of regulatory proteins may affect batteries of downstream genes, but again are indirect and do not necessarily involve substitutions at the loci whose expression changes. Therefore, different approaches have to be taken to identify primary changes in regulatory pathways.

Candidate genes in human evolution

The ultimate goal of microarray analyses, quantitative trait genetics, population genetics or other comparative genetic methods is the identification of genes that are candidates for being causally associated with phenotypic divergence. Although genome-wide, large-scale surveys provide an overview, rigorous tests of causality demand a gene-by-gene approach. In choosing genes to be studied in greater detail, molecular geneticists will be opportunistic, focusing on those loci for which additional information from human or model-animal biology suggests an association with a trait of greater evolutionary interest, such as craniodental development⁷⁸. Therefore, it is unlikely that all traits will be pursued with equal vigour or success.

To implicate a gene in human evolution, two types of data need to be assessed. First, functional evidence that a gene is involved in a developmental, behavioural or physiological trait is required to formulate hypotheses about the role of an individual gene. This may come from analysis of human mutations at a locus (see below). Second, the molecular evolution and population genetics of the locus need to be analysed for evidence of natural selection. Comparison of orthologues from chimpanzees and other primates and mammals, and analyses of intraspecific variation in humans, can reveal signs of positive selection at the sequence level or of a recent 'selective sweep' through a locus (Box 2). Evidence of positive selection has been found at several human loci^{73,79–81}. Although these might be physiologically important (for example, in immunity or reproduction), most genes studied so far are not expected to contribute to the divergence of morphological or behavioural traits. More recently, attention has turned to candidate genes identified from human mutations that do affect such traits.

The evolution of a gene affecting speech. Human medical genetics has made substantial progress, and sophisticated mapping techniques for polymorphisms are accelerating the characterization of genes involved in complex traits, particularly those of medical interest. One of the most provocative reports of late was the identification of the gene *FOXP2* (forkhead box P2), mutations of which are associated with a speech and language disorder⁸². The gene encodes a transcription factor and is therefore expected to control the expression of other genes. The excitement surrounding *FOXP2* stems from the observation that affected individuals appear

to have not an overt impairment, but a lesion in the neural circuitry that affects language processes^{62,82,83}.

Is this a novel human 'language' gene? No, the gene is found in other species. In fact, the human *FOXP2* protein differs from the gorilla and chimpanzee sequence at just two residues, and from the orangutan and mouse sequences at three and four residues, respectively⁷⁶. This history is typical of human and other species' genes, in that most genes have orthologues in other mammals and animals. However, there is the possibility that the two replacements in the *FOXP2* protein that evolved in the human lineage are of functional significance to the origin of language.

To examine whether the *FOXP2* gene has been the target of selection during human evolution, a detailed analysis was undertaken of nucleotide variation over a 14-kilobase (kb) subregion of the large *FOXP2* locus, of amino-acid polymorphism in a segment of the protein, and of chimpanzee and orangutan sequences⁸⁴. An unusual excess was found of rare alleles at the human *FOXP2* locus, and of high-frequency alleles. Reduced genetic variation in neutral linked regions is a predicted consequence of a selective sweep (Box 2), so these observations are consistent with natural selection acting on the *FOXP2* locus. Estimates of the time of fixation of the two amino-acid replacements place them within the last 200,000 yr of human evolution, an intriguing correlation with the estimated age of *H. sapiens*.

However, it should be noted that there are no biological data to support the hypothesis that these amino-acid replacements are functionally important. In the 14-kb region surveyed, more than 100 fixed differences exist; the entire *FOXP2* locus is large (267 kb), and more than 2,000 differences would be expected to exist between the *FOXP2* genes of humans and chimpanzees. No assessment has been made of potential non-coding regulatory sequences that might have contributed to a divergence in the role of *FOXP2* in hominids. Trait differences are often due to changes in regulatory networks that govern development, and need not be in coding regions (although that would be much more convenient, given just two changes in the human *FOXP2* protein). Because *FOXP2* is a transcription factor, changes in *FOXP2* expression could be of functional and evolutionary significance.

The typical genetic architecture that underlies complex traits makes it extremely unlikely that *FOXP2* was the only gene under selection in the evolution of our language capabilities. However, we have no means of assessing the relative contribution of *FOXP2* and other candidate genes. The encouraging lesson of *FOXP2* is that medical genetics has provided an interesting lead into a regulatory network that affects the development of speech ability. Further study of *FOXP2* should lead, at a minimum, to a better understanding of the neurodevelopmental biology of speech and language, and perhaps to more genes with interesting evolutionary histories.

The functions of selected genes

The three genome-scale approaches highlighted here—population genetics, comparative genomics and gene-expression profiling—have all succeeded in finding what each sought: thousands of potential adaptive coding substitutions, regulatory differences in gene expression, and gene duplications and rearrangements. Each has yielded many candidates through which to sift and, interestingly, because of the different search regimens used, there is virtually no overlap in the sets of genetic changes that have been surveyed. It is almost certain that, as in other lineages, all of these types of genetic mechanism have contributed to hominid evolution. The crucial challenge now is to obtain functional data for individual genes and to scrutinize the molecular evolution of candidates for signatures of selection.

To place any candidate gene into a functional context of human trait evolution, advances in primate and human developmental neurobiology will be essential. Non-primates are limited as models

of the development and function of primate and hominid neocortex, and thus as models of the function of proteins such as FOXP2 in the development and elaboration of neural networks. Direct empirical work on developing primates, which faces serious methodological constraints as well as bona fide ethical questions, will be necessary to advance beyond associations and correlations. Testing the functional role of what may be subtle changes in human orthologues of primate genes, a daunting task in the most technically developed model species, will be even more difficult.

There are two immediate avenues to increasing the power of human evolutionary genetics. First, we would increase the value of chimpanzee–human comparative genomics by sequencing the gorilla genome, which is the next earliest branching ape to humans and chimpanzees. This would help us to determine the polarity of genetic changes by distinguishing those changes in the human lineage from those in the chimpanzee lineage. Second, 6×10^9 interbreeding humans is a very large resource for identifying rare mutations (for example, in *FOXP2*) with subtle behavioural or developmental effects, and for mapping genetic variation that underlies morphological variation, both of which could lead to genes that govern the formation of human traits and that might have played a part in hominid evolution.

The fine print below the headlines

It is easy to foresee the media headlines that will announce the completion of the chimpanzee genome. One aim of this article has been to anticipate both the excitement that accomplishment warrants and the more sobering aspects of complex trait genetics and genome-scale evolution. Despite our enhanced understanding of functional genetic architecture, there remains a tendency to associate the development, function or evolution of a trait with single genes (genes ‘for’ speech, cancer and so on). The ghost of ‘hopeful monsters’ still haunts biology and is, unfortunately, a prevalent misconception in the scientific and general press. Perhaps wishful thinking is also an intrinsic part of human nature, but it seems unlikely that the traits that interest us most—bipedalism, skeletal morphology, craniofacial morphology, brain size and speech—were the products of selection of just a few major genes. Just as palaeoanthropology now recognizes a complex pattern of hominin phylogeny and the uncertainties in identifying long-sought common ancestors, and comparative neurobiology now searches for more subtle explanations of human capabilities, the lessons of model-system genetics and comparative genomics should prepare us for the finding that the genetics of hominid trait evolution are, in fact, subtle and complicated.

I underscore this point not just for its scientific relevance, but also because of the larger issues at stake—the meaning of the pursuit of the material basis of human evolution. Evolutionary biology has always faced public resistance. It has been difficult enough to gain acceptance of fundamental ideas using humble finches or fruitflies as examples. We can anticipate even more hostile challenges to human evolutionary genetics. Opponents will be sure to exploit any instances where claims or hypotheses are founded on weak or contradictory data. Witness how the recent scrutiny of data supporting the classic paradigm of industrial melanism has been hijacked by the anti-evolution agenda⁸⁵. The sequencing of the chimpanzee genome will reveal no more directly about the origin of human traits than the sequence of the human genome tells us about how to construct a healthy baby. Headlines may claim more, but we would be well advised to describe this as just the beginning of a large, complex and profoundly important story. □

doi:10.1038/nature01495.

1. Wood, B. Hominid revelations from Chad. *Nature* **418**, 133–135 (2002).
2. Leakey, M. *et al.* New hominin genus from eastern Africa shows diverse middle Pliocene lineages. *Nature* **410**, 433–440 (2001).
3. Brunet, M. *et al.* A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**, 145–151 (2002).

4. Wolpoff, M., Senut, B., Pickford, M. & Hawks, J. Palaeoanthropology: *Sahelanthropus* or ‘*Sahelpithecus*? *Nature* **419**, 581–582 (2002).
5. Groves, C. P. in *Comparative Primate Biology* Vol. 1 (eds Swindler, D. R. & Erwin, J.) 187–218 (Alan R. Liss, New York, 1986).
6. Klein, J. & Takahata, N. *Where Do We Come From? The Molecular Evidence for Human Descent* (Springer, New York, 2002).
7. Wood, B. & Collard, M. The human genus. *Science* **284**, 65–71 (1999).
8. Relethford, J. H. *Genetics and the Search for Modern Human Origins* (Wiley-Liss, New York, 2001).
9. Ruff, C. B., Trinkhaus, E. & Holliday, T. W. Body mass and encephalization in Pleistocene *Homo*. *Nature* **387**, 173–176 (1997).
10. Conroy, G. C. *et al.* Endocranial capacity in an early hominid cranium from Sterkfontein, South Africa. *Science* **280**, 1730–1731 (1998).
11. Conroy, G. C., Weber, G. W., Seidler, H., Recheis, W. & Zur Nedden, E. Endocranial capacity of the Bodo cranium determined from three-dimensional computed tomography. *Am. J. Phys. Anthropol.* **113**, 111–118 (2000).
12. Brodmann, K. Neue Ergebnisse über die vergleichende histologische Lokalisation der Grosshirnrinde mit besonderer Berücksichtigung des Stirnhirns. *Anat. Anz.* **41**, 157–216 (1912).
13. Semendeferi, K., Lu, A., Schenker, N. & Damasio, H. Humans and great apes share a large frontal cortex. *Nature Neurosci.* **5**, 272–276 (2002).
14. Cantalupo, C. & Hopkins, W. D. Asymmetric Broca’s area in great apes. *Nature* **414**, 505 (2001).
15. Hopkins, W. D. & Leavens, D. A. The whole-hand point: The structure and function of pointing from a comparative perspective. *J. Comp. Psychol.* **112**, 95–99 (1998).
16. Holloway, R. L. Indonesian ‘Solo’ (Ngandong) endocranial reconstructions: Some preliminary observations and comparisons with Neandertal and *Homo erectus* groups. *Am. J. Phys. Anthropol.* **53**, 285–295 (1980).
17. Gannon, P. J., Holloway, R. L., Broadfield, D. C. & Braun, A. R. Asymmetry of chimpanzee planum temporale: Humanlike pattern of Wernicke’s brain language area homolog. *Science* **279**, 220–222 (1998).
18. Hopkins, W. D., Marino, L., Rilling, J. K. & MacGregor, L. Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging (MRI). *NeuroReport* **9**, 2913–2918 (1998).
19. Hof, P. R., Nimchinsky, E. A., Perl, D. P. & Erwin, J. M. An unusual population of pyramidal neurons in the anterior cingulate cortex of hominids contains the calcium-binding protein calretinin. *Neurosci. Lett.* **307**, 139–142 (2001).
20. Nimchinsky, E. A. *et al.* A neuronal morphological type unique to humans and great apes. *Proc. Natl Acad. Sci. USA* **96**, 5268–5273 (1999).
21. Buxhoeveden, D., Switala, A., Litaker, M., Roy, E. & Casanova, M. Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex. *Brain Behav. Evol.* **57**, 349–358 (2001).
22. Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K. & Van Hoesen, G. W. Prefrontal cortex in humans and apes: A comparative study of area 10. *Am. J. Phys. Anthropol.* **114**, 224–241 (2001).
23. Penin, X., Berge, C. & Baylac, M. Ontogenetic study of the skull in modern humans and the common chimpanzees: Neotenic hypothesis reconsidered with a tridimensional Procrustes analysis. *Am. J. Phys. Anthropol.* **118**, 50–62 (2002).
24. Ponce de León, M. S. & Zollikofer, C. P. E. Neandertal cranial ontogeny and its implications for late hominid diversity. *Nature* **412**, 534–538 (2001).
25. Lieberman, D. E., McBratney, B. M. & Krovitz, G. The evolution and development of cranial form in *Homo sapiens*. *Proc. Natl Acad. Sci. USA* **99**, 1134–1139 (2002).
26. Williams, F. L., Godfrey, L. R. & Sutherland, M. R. in *Human Evolution through Developmental Change* (eds Minugh-Purvis, N. & McNamara, K. J.) 405–441 (Johns Hopkins Univ. Press, Baltimore, 2002).
27. Dean, C. *et al.* Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins. *Nature* **414**, 628–631 (2001).
28. Rice, S. H. in *Human Evolution through Developmental Change* (eds Minugh-Purvis, N. & McNamara, K. J.) 154–170 (Johns Hopkins Univ. Press, Baltimore, 2002).
29. Gould, S. J. *Ontogeny and Phylogeny* (Belknap, Cambridge, Massachusetts, 1977).
30. Moggi-Cecchi, J. Questions of growth. *Nature* **414**, 595–597 (2001).
31. MacFadden, B. J. Fossil horses from ‘Eohippus’ (*Hyracotherium*) to *Equus*: Scaling, Cope’s Law, and the evolution of body size. *Paleobiology* **12**, 355–369 (1986).
32. Charlesworth, B., Lande, R. & Slatkin, M. A neo-Darwinian commentary on macroevolution. *Evolution* **36**, 474–498 (1982).
33. Schwartz, J. H. Homeobox genes, fossils, and the origin of species. *Anat. Rec.* **257**, 15–31 (1999).
34. Klein, R. G. Archeology and the evolution of human behavior. *Evol. Anthropol.* **9**, 17–36 (2000).
35. Mackay, T. F. C. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–20 (2001).
36. Atchley, W. R., Plummer, A. A. & Riska, B. Genetic analysis of size-scaling patterns in the mouse mandible. *Genetics* **111**, 579–595 (1985).
37. Atchley, W. R. & Zhu, J. Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics* **147**, 765–776 (1997).
38. Doebley, J. & Stec, A. Inheritance of the morphological differences between maize and teosinte: Comparison of results for two F2 populations. *Genetics* **134**, 559–570 (1993).
39. Livshits, G., Roset, A., Yakovenko, K., Trofimov, S. & Kobylansky, E. Genetics of human body size and shape: Body proportions and indices. *Ann. Hum. Biol.* **29**, 271–289 (2002).
40. Brakefield, P. *et al.* Development, plasticity and evolution of butterfly eyespot patterns. *Nature* **384**, 236–242 (1996).
41. True, J. R., Liu, J., Stam, L. F., Zeng, Z. B. & Laurie, C. C. Quantitative genetic analysis of divergence in male secondary sexual traits between *Drosophila simulans* and *Drosophila mauritiana*. *Evolution* **51**, 816–832 (1997).
42. Shubin, N., Wake, D. B. & Crawford, A. J. Morphological variation in the limbs of *Taricha granulosa* (Caudata: Salamandridae): Evolutionary and phylogenetic implications. *Evolution* **49**, 874–884 (1995).
43. Gingerich, P. D. Rates of evolution on the time scale of the evolutionary process. *Genetics* **112/113**, 127–144 (2001).
44. Gingerich, P. D. Rate of evolution: Effects of time and temporal scaling. *Science* **222**, 159–161 (1983).
45. Lauter, N. & Doebley, J. Genetic variation for phenotypically invariant traits detected in teosinte: Implications for the evolution of novel forms. *Genetics* **160**, 333–342 (2002).
46. Polaczyk, P. J., Gasparini, R. & Gibson, G. Naturally occurring genetic variation affects *Drosophila* photoreceptor determination. *Dev. Genes. Evol.* **207**, 462–470 (1998).

47. Rutherford, S. L. & Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **396**, 336–342 (1998).
48. Gibson, G., Wemple, M. & van Helden, S. Potential variance affecting homeotic *Ultrabithorax* and *Antennapedia* phenotypes in *Drosophila melanogaster*. *Genetics* **151**, 1081–1091 (1999).
49. Goldschmidt, R. *The Material Basis of Evolution* (Yale Univ. Press, New Haven, Connecticut, 1940).
50. Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Scientific, Malden, Massachusetts, 2001).
51. Davidson, E. H. *Genomic Regulatory Systems: Development and Evolution* (Academic, San Diego, 2001).
52. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer Associates, Sunderland, Massachusetts, 2002).
53. Laurie, C. C. & Stam, L. F. Molecular dissection of a major gene effect on a quantitative trait: The level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144**, 1559–1564 (1996).
54. Long, A. D., Lyman, R. F., Morgan, A. H., Langley, C. H. & Mackay, T. F. C. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete-scute* complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* **154**, 1255–1269 (2000).
55. Wang, R.-L., Stec, A., Hey, J., Lukens, L. & Doebley, J. The limits of selection during maize domestication. *Nature* **398**, 236–239 (1999).
56. Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017 (1998).
57. Lai, C., Lyman, R. F., Long, A. D., Langley, C. H. & Mackay, T. F. C. Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster*. *Science* **266**, 1697–1702 (1994).
58. Nuzhdin, S. V. & Reiwitich, S. G. Are the same genes responsible for intra- and interspecific variability for sex comb tooth number in *Drosophila*. *Heredity* **84**, 97–102 (2000).
59. Chen, F.-C. & Li, W.-H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **68**, 444–456 (2001).
60. Britten, R. J. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc. Natl Acad. Sci. USA* **99**, 13633–13634 (2002).
61. Fujiyama, A. *et al.* Construction and analysis of a human–chimpanzee comparative clone map. *Science* **295**, 131–134 (2002).
62. Johnson, M. E. *et al.* Positive selection of a gene family during the emergence of humans and African apes. *Nature* **413**, 514–518 (2001).
63. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
64. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1323 (2001).
65. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
66. Li, W.-H. *Molecular Evolution* (Sinauer Associates, Sunderland, Massachusetts, 1997).
67. Kreitman, M. Methods to detect selection in populations with applications to the human. *Ann. Rev. Genomics Hum. Genet.* **1**, 539–559 (2000).
68. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
69. Fay, J. C., Wyckoff, G. J. & Wu, C.-I. Positive and negative selection on the human genome. *Genetics* **158**, 1227–1234 (2001).
70. Ohta, T. Near-neutrality in evolution of genes and gene regulation. *Proc. Natl Acad. Sci. USA* **99**, 16134–16137 (2002).
71. Hughes, A. L. *Adaptive Evolution of Genes and Genomes* (Oxford Univ. Press, New York, 1999).
72. Swanson, W. J. & Vacquier, V. D. The rapid evolution of reproductive proteins. *Nature Rev. Genet.* **3**, 137–144 (2002).
73. Wyckoff, G. J., Wang, W. & Wu, C.-I. Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304–309 (2000).
74. King, M.-C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
75. Nachman, M. W. Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet.* **17**, 481–485 (2001).
76. Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340–343 (2002).
77. Yan, H., Yuan, W., Velculescu, V. E., Vogelstein, B. & Kinzler, K. W. Allelic variation in human gene expression. *Science* **297**, 1143 (2002).
78. McCollum, M. A. & Sharpe, P. T. Developmental genetics and early hominid craniodental evolution. *BioEssays* **23**, 481–493 (2001).
79. Andolfatto, P. Adaptive hitchhiking effects on genome variability. *Curr. Opin. Genet. Dev.* **11**, 635–641 (2001).
80. Diller, K. C., Gilbert, W. A. & Kocher, T. D. Selective sweeps in the human genome: A starting point for identifying genetic differences between modern humans and chimpanzees. *Mol. Biol. Evol.* **19**, 2342–2345 (2002).
81. Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002).
82. Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).
83. Pinker, S. Talk of genetics and vice versa. *Nature* **413**, 465–466 (2001).
84. Enard, W. *et al.* Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* **418**, 869–872 (2002).
85. Coyne, J. A. Evolution under pressure. *Nature* **418**, 19–20 (2002).
86. Kimbel, W. H., Johanson, D. C. & Rak, Y. The first skull and other new discoveries of *Australopithecus afarensis* at Hadar, Ethiopia. *Nature* **368**, 449–451 (1994).
87. Haile-Selassie, Y. Late Miocene hominids from the Middle Awash, Ethiopia. *Nature* **412**, 178–181 (2001).
88. Asfaw, B. *et al.* Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature* **416**, 317–320 (2002).
89. Hacia, J. G. Genome of the apes. *Trends Genet.* **17**, 637–645 (2001).
90. Richmond, B. G., Aiello, L. C. & Wood, B. A. Early hominid limb proportions. *J. Hum. Evol.* **43**, 529–548 (2002).
91. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–596 (1989).
92. Ptak, S. E. & Przeworski, M. Evidence for population growth in humans is confounded by fine-scale population structure. *Trends Genet.* **18**, 1–5 (2002).

Acknowledgements Thanks to B. Hopkins and C. Cantalupo for guidance on Fig. 2, and to L. Olds for illustrations; to B. Williams, A. Kopp, S. Paddock, A. Rokas, D. Bownds, J. Doebley, N. Shubin and J. Crow for comments on the manuscript; to P. N. and J. Carroll for inspiration, and to J. Carroll for preparation of the manuscript. S.B.C. is an Investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to the author (e-mail: sbcarroll@facstaff.wisc.edu).

The genome sequence of the filamentous fungus *Neurospora crassa*

James E. Galagan¹, Sarah E. Calvo¹, Katherine A. Borkovich², Eric U. Selker³, Nick D. Read⁴, David Jaffe¹, William FitzHugh⁵, Li-Jun Ma¹, Serge Smirnov¹, Seth Purcell¹, Bushra Rehman¹, Timothy Elkins¹, Reinhard Engels¹, Shunguang Wang¹, Cydney B. Nielsen¹, Jonathan Butler¹, Matthew Endrizzi¹, Dayong Qui¹, Peter Ianakiev¹, Deborah Bell-Pedersen⁶, Mary Anne Nelson⁷, Margaret Werner-Washburne⁷, Claude P. Selitrennikoff⁸, John A. Kinsey⁹, Edward L. Braun¹⁰, Alex Zelter^{4,11}, Ulrich Schulte¹², Gregory O. Kothe³, Gregory Jedd¹³, Werner Mewes^{14,15}, Chuck Staben¹⁶, Edward Marcotte¹⁷, David Greenberg¹⁸, Alice Roy¹, Karen Foley¹, Jerome Naylor¹, Nicole Stange-Thomann¹, Robert Barrett¹, Sante Gnerre¹, Michael Kamal¹, Manolis Kamvyselis¹, Evan Mauceli¹, Cord Bielke¹⁴, Stephen Rudd¹⁵, Dmitrij Frishman¹⁵, Svetlana Krystofova², Carolyn Rasmussen¹⁹, Robert L. Metzenberg²⁰, David D. Perkins²⁰, Scott Kroken¹⁹, Carlo Cogoni²¹, Giuseppe Macino²¹, David Catcheside²², Weixi Li¹⁶, Robert J. Pratt⁶, Stephen A. Osmani²³, Colin P. C. DeSouza²⁴, Louise Glass¹⁹, Marc J. Orbach²⁵, J. Andrew Berglund³, Rodger Voelker³, Oded Yarden¹¹, Michael Plamann²⁶, Stephan Seiler²⁶, Jay Dunlap²⁷, Alan Radford²⁸, Rodolfo Aramayo⁶, Donald O. Natvig⁷, Lisa A. Alex²⁹, Gertrud Mannhaupt¹⁴, Daniel J. Ebbole³⁰, Michael Freitag³, Ian Paulsen¹⁸, Matthew S. Sachs³¹, Eric S. Lander^{1,32}, Chad Nusbaum¹ & Bruce Birren¹

A list of author affiliations appears at the end of the paper

Neurospora crassa is a central organism in the history of twentieth-century genetics, biochemistry and molecular biology. Here, we report a high-quality draft sequence of the *N. crassa* genome. The approximately 40-megabase genome encodes about 10,000 protein-coding genes—more than twice as many as in the fission yeast *Schizosaccharomyces pombe* and only about 25% fewer than in the fruitfly *Drosophila melanogaster*. Analysis of the gene set yields insights into unexpected aspects of *Neurospora* biology including the identification of genes potentially associated with red light photobiology, genes implicated in secondary metabolism, and important differences in Ca²⁺ signalling as compared with plants and animals. *Neurospora* possesses the widest array of genome defence mechanisms known for any eukaryotic organism, including a process unique to fungi called repeat-induced point mutation (RIP). Genome analysis suggests that RIP has had a profound impact on genome evolution, greatly slowing the creation of new genes through genomic duplication and resulting in a genome with an unusually low proportion of closely related genes.

Research on *Neurospora* in the early part of the twentieth century paved the way for modern genetics and molecular biology. First documented in 1843 as a contaminant of bakeries in Paris¹, *Neurospora* was developed as an experimental organism in the 1920s^{2,3}. Subsequent work on *Neurospora* by Beadle and Tatum⁴ in the 1940s established the relationship between genes and proteins, summarized in the ‘one-gene-one-enzyme’ hypothesis. In the latter half of the century, *Neurospora* had a central role as a model organism, contributing to the fundamental understanding of genome defence systems, DNA methylation, mitochondrial protein import, circadian rhythms, post-transcriptional gene silencing and DNA repair⁵. Because *Neurospora* is a multicellular filamentous fungus, it has also provided a system to study cellular differentiation and development as well as other aspects of eukaryotic biology⁶.

The legacy of over 70 years of research⁷, coupled with the availability of molecular and genetic tools, offers enormous potential for continued discovery. The sequencing of the *N. crassa* genome was undertaken to maximize this potential. Here, we report an initial sequence and analysis of the *Neurospora* genome.

Neurospora genome sequence

The *Neurospora* genome is much larger (greater than 40 megabases (Mb)) than that of *S. pombe* and *Saccharomyces cerevisiae* (both about 12 Mb). Accordingly, first we sought to produce and analyse a high-quality draft sequence *en route* to a finished sequence.

The genome sequence was assembled from deep whole-genome shotgun (WGS) coverage obtained by paired-end sequencing from a variety of clone types (Supplementary Information). In all, the data provided an average of >20-fold sequence coverage and >98-fold physical coverage of the genome. The Arachne package⁸ was used to assemble the draft genome sequence. The resulting assembly

consists of 958 sequence contigs with a total length of 38.6 Mb (Table 1) and an N50 length of 114.5 kilobases (kb) (that is, 50% of all bases are contained in contigs of at least 114.5 kb). Contigs were assembled into 163 scaffolds with a total length of 39.9 Mb (including gaps between contigs) and an N50 length of 1.56 Mb.

Most of the assembly (97%) is contained in the 44 largest scaffolds, and there are 38 tiny scaffolds with lengths <4 kb. Forty-two of the large scaffolds (and one of the smaller ones) could be anchored readily to the *Neurospora* genetic map⁷ by virtue of their containing genetic markers with sequence.

The assembly has long-range continuity, with the N50 scaffold size being nearly 1,000-fold larger than the average gene size. The assembly represents the vast majority of the genome, as assessed by comparison with available finished sequence and genetic markers. It contains 99.13% of available finished sequence (17 Mb from linkage groups II and V⁹) and all of the 252 genetic markers with sequence. This estimate, however, does not account for unusual genomic regions such as the ribosomal DNA repeats, centromeres and telomeres; such regions may contain about 1.7 Mb of additional sequence¹⁰, corresponding to 2–3% of the genome that cannot be assembled readily with available techniques. The long-range continuity of the assembly was also confirmed by comparisons with previously described bacterial artificial chromosome (BAC) physical maps for linkage groups II and V¹¹, as only one discrepancy was noted.

The assembly also has high accuracy, with 99.5% of the sequence having Arachne quality scores ≥30. Comparison with the 17 Mb of finished sequence confirms the sequence accuracy, with a discrepancy rate for this subset of less than 10^{−5}. The comparison also largely confirms the assembly, as only 12 minor discrepancies were identified (Supplementary Information).

Genes

Gene count and basic characteristics

A total of 10,082 protein-coding genes (9,200 longer than 100 amino acids) were predicted (Table 1 and Supplementary S0). This constitutes nearly twice as many genes as in *S. pombe* (about 4,800) and *S. cerevisiae* (about 6,300), and nearly as many as in *D. melanogaster* (about 14,300). Genes cover at least 44% of the genome sequence with an average gene density of one gene per 3.7 kb. The average gene length of 1.67 kb is slightly longer than the 1.4-kb average gene length for both *S. cerevisiae* and *S. pombe*. The difference in gene length is due to the greater number of introns in *Neurospora* genes—an average of 1.7 introns per gene with an average intron size of 134 nucleotides. Notably, most predicted *Neurospora* introns lack a polypyrimidine tract, which is common in other eukaryotic introns, but do contain a strong branchpoint sequence (Supplementary Information).

Comparative analysis

A total of 4,140 (41%) *Neurospora* proteins lack significant matches to known proteins from public databases (Table 1), reflecting the early stage of fungal genome exploration and the diversity of fungal genes remaining to be described. Furthermore, 5,805 (57%) *Neurospora* proteins do not have significant matches to genes in either of the sequenced yeast species (Supplementary Information). When compared to sequenced eukaryotes, a total of 1,421 (14%) *Neurospora* genes display best BLASTP matches to proteins in either plants or animals (Supplementary Information). Of these, 584 lack high-scoring hits to either sequenced yeast species. These data reflect the biology shared by filamentous fungi and higher eukaryotes, which in a number of cases is absent in the yeasts.

Epigenetics, genome defence and genome evolution

Neurospora is an important model for the study of epigenetic phenomena, possessing a wide variety of epigenetic mechanisms and related genome defence mechanisms. The most remarkable of these mechanisms is repeat-induced point mutation (RIP), a process unique to fungi.

Repeat-induced point mutation

First discovered in *Neurospora*^{12,13}, RIP is a process that efficiently detects and mutates both copies of a sequence duplication. RIP acts during the haploid dikaryotic stage of the *Neurospora* sexual reproductive cycle, causing numerous C•G to T•A mutations within duplicated sequences. In a single passage through the sexual cycle, up to 30% of the C•G pairs in duplicated sequences can be mutated, with a strong preference for C to T mutations occurring at CpA dinucleotides¹⁴. The pattern of mutations produces a

characteristic skewing of dinucleotide frequencies that allows RIP-mutated sequences to be detected accurately¹⁵. RIP requires a minimal duplicated sequence length of about 400 base pairs (bp)¹⁶ and greater than roughly 80% sequence identity between duplicates¹⁷. In addition to suffering mutations, RIP-mutated sequences are frequent targets for DNA methylation. As with mammals, DNA methylation has been shown to cause gene silencing in *Neurospora*¹⁸. RIP thus mutates and epigenetically silences repetitive DNA.

RIP has been proposed to act as a defence against selfish or mobile DNA¹³. However, because RIP mutation and methylation can extend beyond the bounds of duplicated sequences¹⁹, RIP can have both mutational and epigenetic effects on neighbouring unique sequences. Furthermore, RIP acts on all duplicated sequences, including those arising from large-scale chromosomal duplications as well as gene duplications²⁰. The presence of RIP thus has profound consequences for the evolution of the *Neurospora* genome. Indeed, it has been proposed that RIP might prevent gene innovation through gene duplication^{13,21}. With the availability of the *Neurospora* genome sequence, we were able to address this hypothesis.

Multigene families

To investigate the impact of RIP on protein families in *Neurospora*, genes were clustered into 'multigene families' on the basis of an all versus all comparison of protein sequences (see Methods). As shown in Fig. 1, the percentage of genes in multigene families in selected sequenced eukaryotes is correlated with genome size. However, in marked contrast to the other analysed organisms, *Neurospora* possesses many fewer genes in multigene families than expected. When the analysis is expanded to include an additional 17 sequenced prokaryotes (Supplementary Information), only *Mycoplasma genitalium*, *Mycoplasma pulminus*, *Ureaplasma urealyticum* and *Vibrio cholerae* display a correspondingly small proportion of genes in families. This is noteworthy considering that the *Mycoplasma* genus is thought to have undergone reductive evolution and represent minimal life forms²².

Our analysis reveals another characteristic of *Neurospora* gene families. Unlike other sequenced eukaryotes, *Neurospora* possesses only a handful of highly similar gene pairs. Figure 2 displays histograms of amino acid and nucleotide similarities between each gene in the six organisms analysed and the best-matching gene in that organism. A significant proportion of genes have best matches with greater than 80% amino acid and nucleotide identity in all the organisms considered except *Neurospora*. *Neurospora* contains only eight genes with top matches of greater than 80% amino acid or coding sequence identity. This value is significant because, as described above, RIP mutates duplicated sequences that display greater than about 80% nucleotide similarity. Thus, the small proportion of genes in multigene families and the near absence of highly similar genes are consistent with the actions of RIP.

An example of the lack of highly similar genes in multigene families is revealed in an analysis of predicted major facilitator superfamily (MFS) sugar transporters (Fig. 3). *Neurospora* has about the same number of predicted MFS sugar transporters as *S. cerevisiae*. However, a phylogenetic analysis of fungal sugar transporters indicates that the *Neurospora* proteins are substantially more divergent than those of *S. cerevisiae* as well as those of *S. pombe*. Furthermore, the *Neurospora* transporters contain no apparent instances of recent duplication. In contrast, most of the *S. cerevisiae* HXT hexose and *S. pombe* GHT transporters represent two relatively recent and independent expansions and include very recently duplicated genes. Thus, despite a diversity of MFS sugar transporters, *Neurospora* seems to lack close paralogues in this gene family, consistent with the results of the genome-wide multigene family analysis.

Table 1 *Neurospora crassa* genome features

Feature	Value
General	
Size (bp) (assembly 5)	38,639,769
Chromosomes	7
G + C content (%)	50
Protein-coding genes	10,082
Protein-coding genes >100 amino acids	9,200
tRNA genes	424
5S rRNA genes	74
Per cent coding	44
Per cent intronic	6
Average gene size (bp)	1,673 (481 amino acids)
Average intergenic distance (bp)	1,953
Predicted protein-coding sequences	
Identified by similarity to known sequences	1,336 (13%)
Conserved hypothetical proteins	4,606 (46%)
Predicted proteins (no similarity to known sequences)	4,140 (41%)

Analyses of other gene families yielded similar results (data not shown). Furthermore, the paucity of closely related sequences is evident not only at the level of complete genes, but even at the level of individual exons, protein domains and protein architectures (Supplementary S4).

Gene evolution through gene duplication

The above results suggest that RIP has had a powerful impact in suppressing the creation of new genes or partial genes through genomic duplication. This is consistent with the large number of mutations induced in duplicated sequences by RIP. Computer simulations (see Methods) indicate that after a gene duplication, each copy has an 80% probability of acquiring an in-frame stop codon after only a single round of RIP and a 99.5% probability by the point that RIP has mutated the copies to less than 85% nucleotide similarity. The high frequency of stop codons reflects the preference of RIP for mutating CpA to TpA, increasing the prevalence of the stop codons TAA and TAG.

These results raise the critical question of whether any significant gene duplication has occurred in *Neurospora* subsequent to the acquisition of RIP. We searched for empirical evidence of duplicated genes that have survived RIP by analysing the set of *Neurospora* coding sequences using two different measures¹⁵ for detecting RIP-mutated sequences (see Methods). These measures use the characteristic skewing of dinucleotides produced by RIP to detect mutated sequences. According to these measures, only 59 of the 9,200 predicted genes encoding proteins ≥ 100 amino acids show evidence of mutation by RIP. Of these, only eight consist of pairs of predicted duplicated genes (genes in the same multigene family) in which both copies are predicted to be RIP-mutated. Thus, few pairs of duplicated genes display evidence of having both survived RIP (Supplementary Information).

Gene duplication is thought to have a primary role in the innovation of new genes²³. However, taken together, our data support the conclusion that most, if not all, paralogous genes in *Neurospora* duplicated and diverged before the emergence of RIP, and since that point the evolution of new genes through gene duplication has been virtually arrested. This conclusion raises the question of whether and how *Neurospora* is able to evolve new genes. A number of mechanisms that do not involve gene duplication are conceivable, although ultimately a conclusive analysis may be possible only by comparing the genome of *Neurospora* with the genomes of closely related species to illuminate recent evolutionary

history. Nonetheless, our results indicate that the cost to *Neurospora* of increased genome security through RIP is a significant impact on the evolution of new gene functions through gene duplication.

Repetitive DNA

An analysis of repeat sequences longer than 200 bp and with greater than 65% similarity (see Methods) revealed that 10% of the *Neurospora* assembly consists of repeat sequences, consistent with previously reported estimates²¹.

The repeat sequence of *Neurospora* provides a testament to the efficiency of RIP. Applying the measures of RIP mentioned above to the *Neurospora* genome revealed that most of the repetitive sequences (81%) in *Neurospora* have been mutated by RIP. Conversely, only 18% of predicted RIP-mutated sequence is non-repetitive, potentially reflecting loss of the corresponding duplicated sequence. As described above, duplications greater than about 400 bp are susceptible to RIP¹⁶. In keeping with this, we observe that over 97% of genomic repeats greater than 400 bp in length are RIP-mutated. Moreover, repeats longer than 400 bp clustered by sequence similarity display an average sequence identity within clusters of 78%, with 93% of clusters displaying an average identity of less than 85%. This corresponds to previous estimates indicating

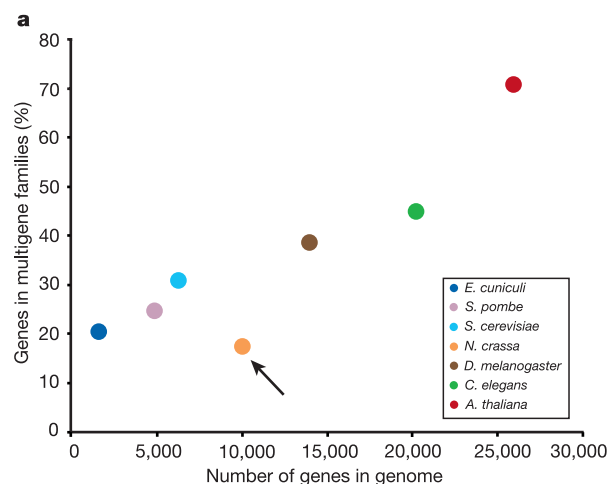


Figure 1 *Neurospora* has a low proportion of genes in multigene families. The graph displays the proportion of genes in multigene families (see Methods) as a function of the number of genes in the genomes of selected sequenced eukaryotic organisms. The arrow indicates *Neurospora*. See text for more details.

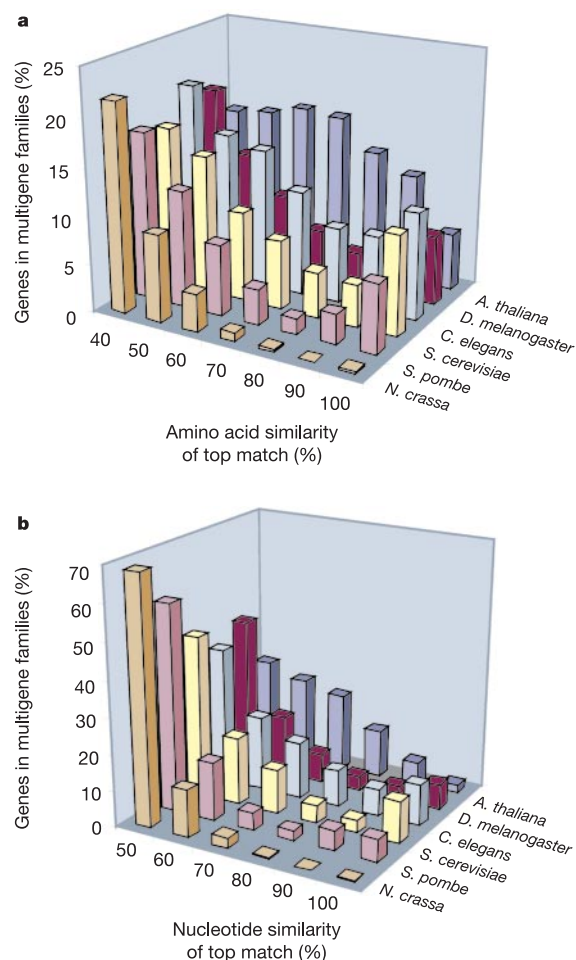


Figure 2 *Neurospora* possesses few highly similar genes. **a**, **b**, Histogram of amino acid (**a**) and nucleotide (**b**) per cent identity of top-scoring self-matches for genes in selected sequenced eukaryotic genomes. For each organism, the protein and coding regions for each gene (not including pseudogenes) were compared to those of every other gene in the same genome using BLASTX. Top-scoring matches were aligned using ClustalW and per cent identities calculated. In contrast to other eukaryotes, *Neurospora* possesses only eight genes with a top match of greater than 80% amino acid or nucleotide identity.

that RIP requires greater than about 80% sequence identity to detect duplicated sequences.

Consistent with the hypothesis that RIP acts as a defence mechanism against selfish DNA¹³, no intact mobile elements were identified. Furthermore, a significant proportion of the *Neurospora* RIP-mutated sequence (46% of repetitive nucleotides) can be identified as relics of mobile elements (Supplementary Information).

Ribosomal RNA

The only large repetitive sequences known to have survived RIP in *Neurospora* are the approximately 175–200 copies¹⁰ of the large rDNA tandem repeat containing the 17S, 5.8S and 25S rRNA genes. As in higher eukaryotes, these tandem repeats occur within the nucleolar organizer region (NOR), and their resistance to RIP seems to stem from this localization¹³. Within the genome sequence we found several copies of the rDNA repeat outside the NOR. In every

case, they display evidence of mutation by RIP, consistent with previous observations¹³. Thus, the sequence of the rDNA repeat does not in itself seem to confer resistance to RIP.

The 5S rRNA genes in *Neurospora* have survived RIP in a different manner. In contrast to most higher eukaryotes in which the 5S rRNA genes form tandem repeats, the 5S genes are dispersed throughout the genome in *Neurospora*²⁴. A total of 74 copies comprising several different subtypes of 5S rDNA are dispersed through all seven chromosomes. This dispersal coupled with their small size (approximately 120 nucleotides) ensures that they are not recognized by RIP.

DNA methylation

Neurospora has been used extensively as a model for studying DNA methylation in eukaryotes²⁵. The *Neurospora* genome includes two potential cytosine DNA methyltransferase genes. One, called *dim-2*, is required for all known DNA methylation²⁶. The other, called *rid*, is

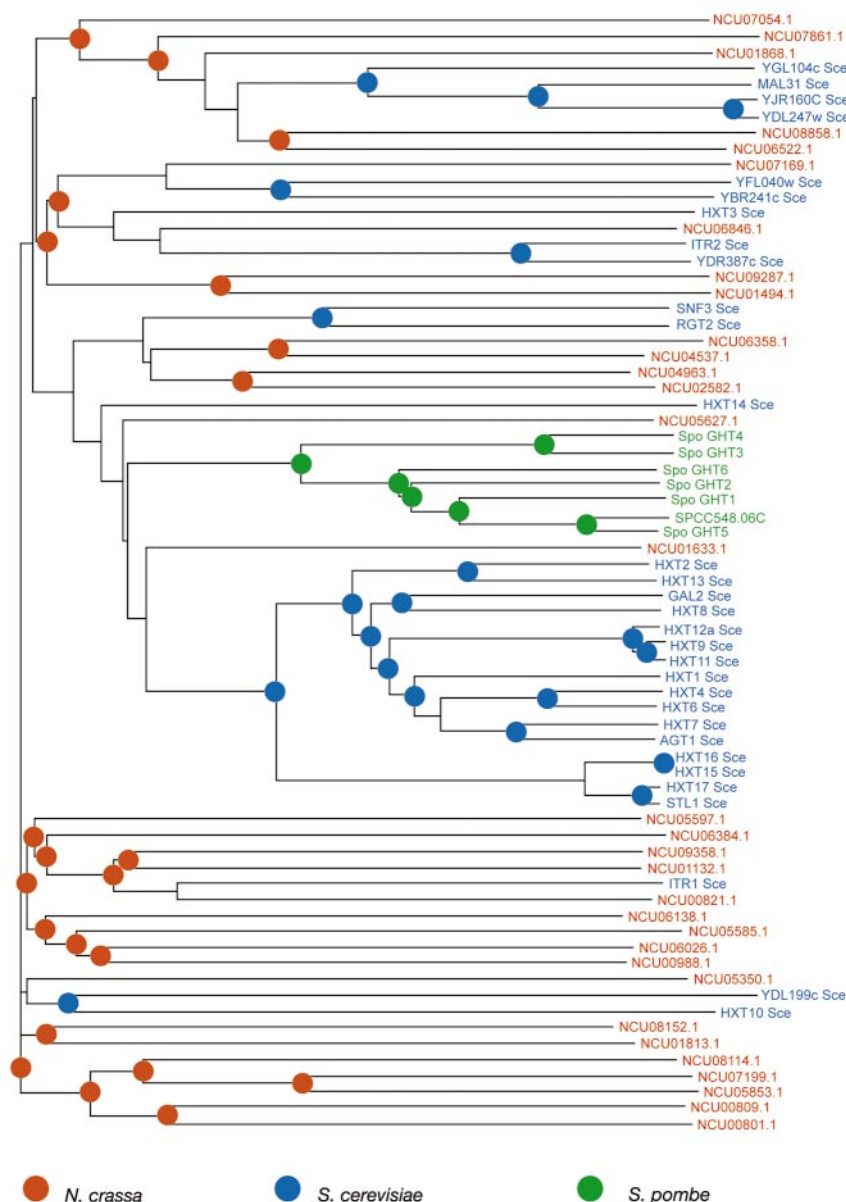


Figure 3 Example of lack of recent duplications in a *Neurospora* gene family. Phylogenetic tree of major facilitator superfamily (MFS) sugar transporters from *S. cerevisiae*, *S. pombe* and *Neurospora*. Coloured dots represent branching points between predicted paralogous

genes in *Neurospora* (red), *S. cerevisiae* (blue) and *S. pombe* (green). In contrast to both yeast species, *Neurospora* transporters contain no predicted instances of recent paralogous duplication.

required for RIP and is a member of a family found thus far only in filamentous fungi²⁷. In *Neurospora*, an estimated 1.5% of cytosines are methylated^{28,29}, and it has been suggested that nearly all DNA methylation is a result of RIP^{1,30,31}.

Plasmid reads for *Neurospora* were sequenced from libraries cloned separately in methylation-tolerant and methylation-intolerant strains of *Escherichia coli*. Although not intended for this purpose, these libraries provided a basis for predicting DNA methylation by comparing the representation of regions in sequence obtained from each library (see Methods). Testing the accuracy of such predictions, we found that 8 of 10 regions predicted to be methylated were experimentally confirmed as such. The predictions thus have good specificity—although they lack sensitivity (see Methods).

The specificity of the predictions provides insight into the pattern of methylation in the *Neurospora* genome. Regions predicted to be methylated show a marked correspondence to regions predicted to be repetitive and RIP-mutated (Fig. 4). Fully 85% correspond to predicted RIP-mutated sequences. However, a small proportion (10%) corresponds to predicted non-repetitive and non-RIP-mutated sequence. In two out of ten such cases, both the methylation and the non-repetitive nature of these sequences were experimentally verified. This raises the possibility that methylation in *Neurospora* may also have non-defence roles, as proposed for higher organisms.

RNA silencing

Post-transcriptional gene silencing (PTGS), or RNA silencing, is widespread among organisms and is increasingly being recognized as a principal switch for controlling eukaryotic gene expression³². RNA-silencing pathways are thought to be derived from ancestral natural defence systems directed against invading nucleic acids³³.

Consistent with this, all known PTGS mechanisms share similar components³⁴.

Neurospora possesses two RNA-silencing pathways. The first, called quelling, acts during vegetative growth. This pathway was uncovered through the study of three genes, *qde-1*, *qde-2* and *qde-3*, coding respectively for an RNA-dependent RNA polymerase (RdRP), an argonaute and a RecQ helicase³⁵. The second pathway, called meiotic silencing, acts during sexual reproduction^{36,37}. Before our analysis, a gene called *sad-1*, encoding an RdRP, had been identified for this pathway³⁸.

Our analysis of the *Neurospora* genome sequence uncovered several additional genes implicated in RNA silencing (Table 2). These include one RdRP, one argonaute-like protein and one RecQ-like helicase, as well as two dicer-like ribonucleases. A phylogenetic analysis (Supplementary S7) of the predicted RdRPs, argonaute-like proteins and dicer-like proteins indicates that the *Neurospora* genes comprise two paralogous sets. One set includes the three *qde* genes and is thus predicted to correspond to the quelling pathway. The other set includes *sad-1*, and in phylogenetic trees these genes branch consistently with those of the single pathway observed in *S. pombe*^{37,39}. On the basis of this analysis, we predict that one of the identified dicers, *Sms-3*, belongs phylogenetically to the meiotic silencing pathway, whereas the other, *dcl-2*, belongs to the quelling pathway (Table 2). In addition, we predicted that the identified argonaute, *Sms-2*, also belongs phylogenetically to the meiotic silencing pathway. Subsequent experimental work has supported roles for *Sms-2* (ref. 40) and *Sms-3* (M. McLaughlin, D. W. Lee, R. Pratt and R. Aramayo, manuscript in preparation) in meiotic silencing. Taken together, these results suggest that meiotic silencing and quelling represent two phylogenetically distinct RNA-dependent silencing pathways. We further hypothesize that both might have evolved from a single ancestral RNA-silencing pathway.

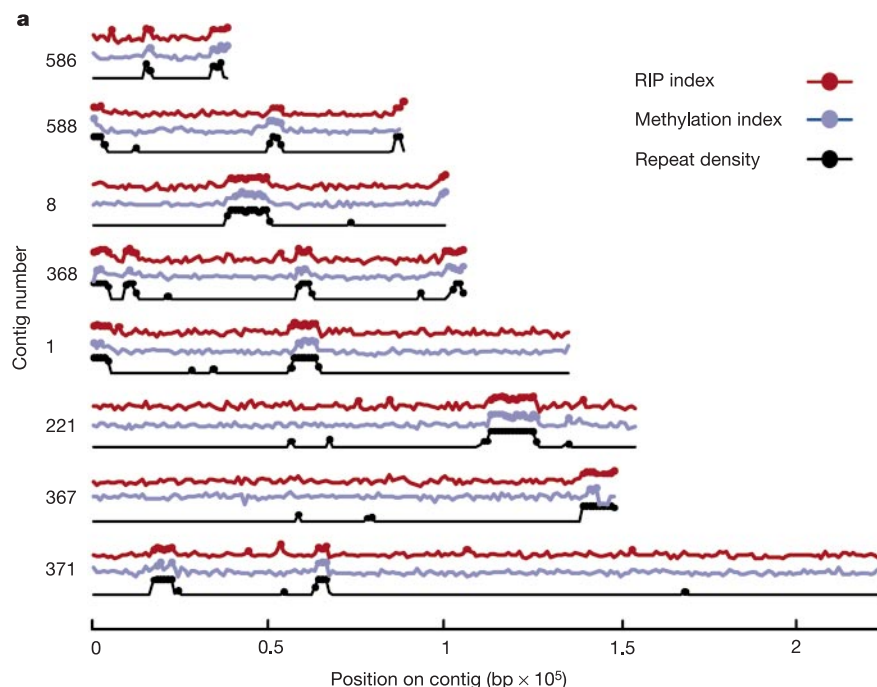


Figure 4 Correspondence between predicted RIP, methylation and repetitive DNA. Prediction of RIP, methylation and repeat sequence in 1-kb windows for selected contigs. Red lines plot the TpA/ApT RIP index (see Methods); red dots indicate windows predicted to be RIP-mutated (TpA/ApT > 1.2). Blue lines plot the proportion of reads from the methyl-tolerant library; blue dots indicate windows predicted to be methylated based on

>70% methyl-tolerant reads (see Methods). Black lines plot repeat content as a fraction of nucleotides in each window that is in repetitive sequence; black dots indicate windows with >50% repeat sequence. Contigs were selected to illustrate regions predicted as methylated.

Fungal biology and evolution

The *Neurospora* genome sequence provides an opportunity to study the genetic basis underlying the extraordinary biochemical and metabolic diversity exhibited by a filamentous fungus. Our analysis of the genome sequence has resulted in a number of surprising insights into the biology and evolution of *Neurospora* and other filamentous fungi.

Cell signalling and environmental responses

Discovery of putative red-light-sensing genes

Blue light is an important regulator of *Neurospora* growth and development, affecting the circadian rhythm of conidiation, carotenogenesis of hyphae and numerous facets of sexual development⁴¹. Although *Neurospora* photobiology has been studied intensively for more than two decades, the genome sequence has nonetheless revealed a number of previously uncharacterized sequences with similarity to blue-light-sensing genes, including both a cryptochrome homologue and a gene whose product contains a single PAS/LOV-type domain associated with light sensing.

Furthermore, *Neurospora* possesses two putative phytochrome homologues most similar to bacteriophytochromes—genes known for their role in red light sensing in prokaryotes—and a putative homologue of the *Aspergillus nidulans* velvet gene implicated in the regulation of both red and blue light responses. The presence of these genes is unexpected given that no red light photobiology has been described for *Neurospora* so far. It has been shown recently that in addition to red light sensing, some *Arabidopsis* phytochromes associate with cryptochromes to have a role in blue light sensing and signalling⁴². Therefore, the two phytochromes and the velvet homologue may also regulate this aspect of *Neurospora* photobiology.

Importance of two-component signalling in filamentous fungi

Mitogen-activated protein kinase (MAPK) pathways integrate signals from multiple receptor pathways including two-component signalling systems⁴³. The basic two-component system consists of a histidine kinase and a cognate response regulator. The nine MAPK pathway proteins identified in the *Neurospora* genome sequence (Fig. 5) correspond to those found in *S. pombe* and *S. cerevisiae*, indicating that the basic MAPK machinery is conserved between these species. In contrast, *Neurospora* has a significantly expanded complement of 11 histidine kinases, as compared with one in *S. cerevisiae* and three in *S. pombe*. Two of the 11 genes have been characterized previously in *Neurospora*⁴⁴, whereas a third is similar to proteins in *Aspergillus fumigatus* and *A. nidulans* that affect conidiation (L. A. Alex and M. I. Simon, unpublished observations; see also ref. 44). Functions for the remaining genes are unknown, although seven (including the two phytochromes discussed above) contain PAS/PAC domains, implicating them in oxygen and light

responses. This number of histidine kinases suggests a larger role than previously expected, and reveals filamentous fungi to be more similar in this regard to plants, where two-component systems are abundant, than to animals, where these systems are absent.

A new family of G-protein-coupled receptors

Eukaryotic cells sense many environmental stimuli through seven-transmembrane-helix, G-protein-coupled receptors (GPCRs)⁴⁵. Our analysis indicates that *Neurospora* possesses ten predicted seven-transmembrane-helix proteins (Fig. 5), three of which belong to a new class not previously identified in any fungus. These three genes encode proteins similar to cyclic AMP GPCRs from the protists *Dictyostelium discoideum*⁴⁶ and *Polysphondylium pallidum*, and also to predicted proteins from *Arabidopsis thaliana*⁴⁷ and *Caenorhabditis elegans*. The *D. discoideum* proteins sense cAMP levels during chemotaxis and multicellular development⁴⁸. This suggests a possible analogous function in *Neurospora*. The existence of an extracellular cAMP signalling pathway has never been demonstrated previously in any fungal system.

In support of this hypothesis, along with the presence of putative cAMP receptors, *Neurospora* was found to possess the full complement of proteins required for the synthesis and degradation of cAMP. Furthermore, *Neurospora* wild-type strains accumulate cAMP in the extracellular medium⁴⁹, although a role in extracellular signalling has not been established. Taken together, these data suggest the possibility that cAMP or a related molecule may serve as an extracellular signal in *Neurospora*.

Ca²⁺ sensory transduction in filamentous fungi

A considerable body of evidence, primarily from pharmacological studies, indicates that Ca²⁺ signalling regulates numerous processes in filamentous fungi⁵⁰. However, the identification of the main components of even one Ca²⁺-mediated response pathway in filamentous fungi has remained elusive. The genome sequence of *Neurospora* has provided over 25 of the proteins likely to be necessary for Ca²⁺ signalling in filamentous fungi (Fig. 5).

A notable difference between Ca²⁺ signalling in *Neurospora* as compared with plants and animals was revealed by the genome sequence. An important aspect of Ca²⁺ signalling in plant and animal cells involves Ca²⁺ release from internal stores. This is commonly mediated by the second messengers inositol-1,4,5-trisphosphate (InsP₃) and cADP ribose, or by Ca²⁺-induced Ca²⁺ release⁵¹. InsP₃ is present within *Neurospora* hyphae⁵², and physiological evidence including intracellular membrane-associated, InsP₃-activated Ca²⁺ channel activity supports a role in Ca²⁺ signalling^{53,54}. In spite of this, *Neurospora* (and *S. cerevisiae*) lacks recognizable InsP₃ receptors. In addition, neither ADP ribosyl cyclase nor ryanodine receptor proteins, principal components of Ca²⁺ release mechanisms in plant and animal cells, are found in *Neurospora*. These observations raise the question of whether other

Table 2 *Neurospora* has two RNA-silencing pathways

Predicted protein	<i>Neurospora</i>	<i>A. fumigatus</i> *	<i>S. pombe</i> †	Pathway‡
RNA-directed RNA polymerases	<i>qde-1</i> (NCU07534.1) <i>sad-1</i> (NCU02178.1) <i>rpp-3</i> (NCU08435.1)	<i>rppA</i> (contig 158) <i>rppB</i> (contig 472) —	— <i>rdp1</i> ⁺ (SPAC6F12.09) —	Quelling Meiotic silencing Unknown
Argonaute-like, related to translation initiation factors	<i>qde-2</i> (NCU04730.1) <i>Sms-2</i> (NCU09434.1)	<i>ppdA</i> (contig 720) <i>ppdB</i> (contig 196)	— <i>ago1</i> ⁺ (SPCC736.11)	Quelling Meiotic silencing
Dicer-like, related to SFII-RNase III ribonucleases of the carpel factory	<i>dcl-2</i> (NCU06766.1) <i>Sms-3</i> (NCU08270.1)	<i>dclB</i> (contig 618) <i>dclA</i> (contig 310)	— <i>dcr1</i> ⁺ (SPCC584.10C)	Quelling Meiotic silencing
RecQ helicase-like, related to Bloom's and Werner syndrome helicases	<i>qde-3</i> (NCU08598.1) <i>RecQ-2</i> (NCU03337.1) [‡]	<i>rqhA</i> (contig 443)§ <i>rqhB</i> (contig 58)§	— <i>hus2</i> ⁺ (SPAC2G11.12)	Quelling Unknown

* Unfinished *A. fumigatus* genome project (<http://www.tigr.org>).

† *Schizosaccharomyces pombe* genome project (<http://www.genedb.org>).

‡ Pathway assigned on the basis of either known experimental data for *qde-1*, *qde-2* and *qde-3* (quelling pathway); or *sad-1*, *Sms-2* and *Sms-3* (meiotic silencing pathway); or predicted on the basis of phylogenetic analysis.

§ RecQ helicase-like (*rqh*).

second messenger systems responsible for Ca^{2+} release from internal stores remain to be discovered in filamentous fungi.

Growth and development

Hyphal growth

True hyphae produced by filamentous fungi are tubular structures consisting of cellular compartments that are usually delineated by incomplete septa⁵. In contrast, the pseudohyphae produced by yeasts consist of chains of uninucleate elongated cells⁵⁵ with no apparent cytoplasmic continuity. The molecular mechanisms underlying these two modes of growth are not well understood.

The two signalling pathways that regulate pseudohyphal growth in *S. cerevisiae*—the MAPK and cAMP modules—are both conserved in the *Neurospora* genome. In *Candida albicans*, capable of pseudohyphal, true hyphal and budding growth, both pathways are required for true hyphal production, suggesting a similar role in *Neurospora*⁵⁶. However, at least three transcription factors—Tec1p, Flo8p and Sfl1p—specifically required for regulating pseudohyphal development in *S. cerevisiae*⁵⁶ were not found in *Neurospora*. Conversely, *Neurospora* possesses a gene with similarity to a transcription factor necessary for hyphal growth in *C. albicans*³⁶ (Efg1). This transcription factor is not required for pseudohyphal growth in *C. albicans*, nor is the homologous protein in *S. cerevisiae* (Phd1p)⁵⁶. More study of the complex pathways underlying these modes of growth is required. Nonetheless, these data clarify in part the distinctions and similarities between the signalling pathways and regulatory components of hyphal and pseudohyphal growth.

The macroconidiation pathway differs from that in *A. nidulans*

Macroconidia are asexual spores common to filamentous fungi but absent from yeast^{5,57}. Components of the macroconidiation path-

way have been identified in both *Neurospora* and the filamentous fungus *A. nidulans*, and known upstream signalling proteins seem to be conserved in both species⁵⁸. In contrast, there is little conservation of downstream components between the two fungi. In *Neurospora*, the *acon-2*, *acon-3*, *fld* and *fl* genes are essential for conidiation⁵, whereas in *A. nidulans*, the FlbC, FlbD, BrIA, AbaA and WetA gene products are required. Our analysis of the genome sequence revealed that *Neurospora* possesses no FlbC, BrIA or AbaA homologues, and a protein with only very weak similarity to approximately 100 amino acids at the carboxy terminus of WetA. These data make clear that the molecular machinery underlying macroconidiation in *Neurospora* differs significantly from that in *A. nidulans*.

Secondary metabolism

The fungal kingdom produces a vast array of small, bioactive compounds termed secondary metabolites that are best known for their roles as pigments, antibiotics and mycotoxins. With the exception of carotenoid and melanin pigment synthesis, *Neurospora* has not been shown to possess secondary metabolism. It was thus surprising that the *Neurospora* genome sequence revealed a number of putative genes for secondary metabolite production.

Non-ribosomal peptide synthetases

Three predicted non-ribosomal peptide synthetase (NRPS) genes and one NRPS-related gene were identified in the *Neurospora* genome sequence (Fig. 6). By phylogenomic analysis, one NRPS gene is orthologous to an *Aureobasidium pullulans* NRPS. The most closely related NRPS of known function is *sid2* of *Ustilago maydis*, which is responsible for production of hydroxamate siderophores⁵⁹. The remaining two are of unknown function, although one is orthologous to an NRPS in *Magnaporthe grisea*, and the other is

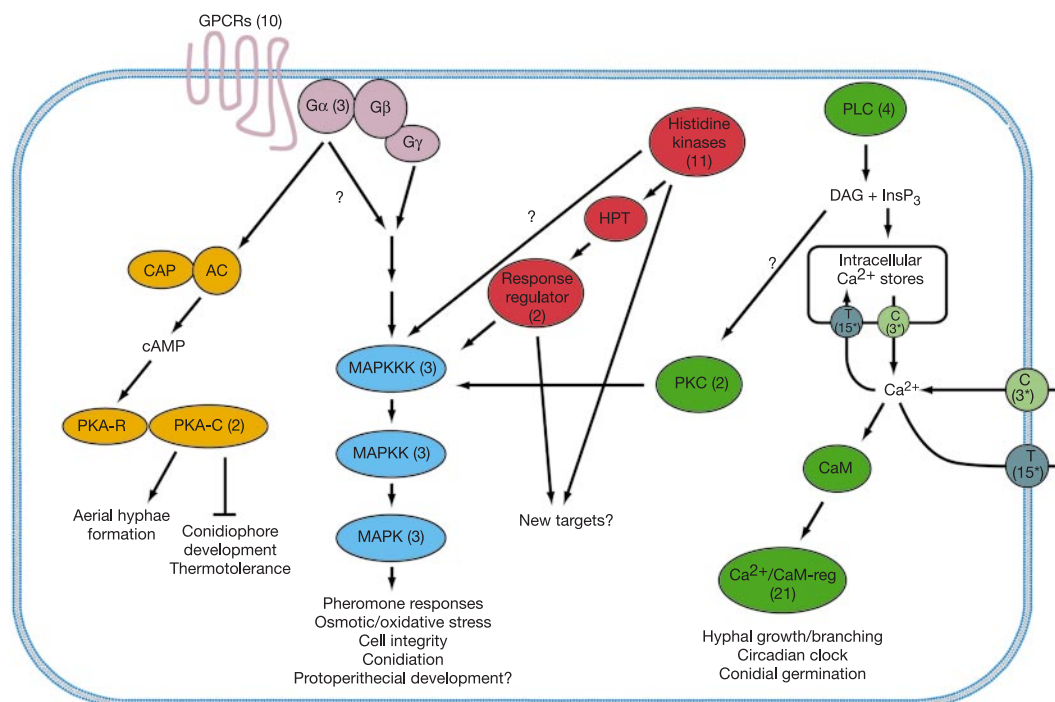


Figure 5 Overview of major intracellular signalling pathways in *Neurospora*. The numbers identified for each gene are in parentheses. An asterisk indicates that the location in the plasma membrane and/or organelle membranes is not determined. AC, adenylyl cyclase; C, Ca^{2+} channel protein; CaM, calmodulin; Ca^{2+} /CaM-reg, calcium- and calmodulin-regulated protein; CAP, cyclase-associated protein; DAG diacylglycerol; GPCR, G-protein-coupled receptor; Gα, G protein α-subunit; Gβ, G protein β-subunit; Gγ, G protein

γ-subunit; HPT, histidine-containing phosphotransfer domain protein; MAPK, MAP kinase; MAPKK, MAPK kinase; MAPKKK, MAPKK kinase; PKA-C, protein kinase A catalytic subunit; PKA-R, protein kinase A regulatory subunit; PLC, phospholipase C; PKC, protein kinase C; T, Ca^{2+} transport protein (P-type Ca^{2+} ATPase, $\text{H}^{+}/\text{Ca}^{2+}$ exchanger, or $\text{Na}^{+}/\text{Ca}^{2+}$ exchanger).

orthologous to an NRPS found in all other filamentous ascomycetes with genome sequence (see Methods). The NRPS-related gene shares 66% amino acid identity with the *CPS1* gene product that contributes to the virulence of *Cochliobolus heterostrophus*, *C. victoriae* and *Gibberella zeae*⁶⁰.

Polyketide synthases

Seven polyketide synthase (PKS) genes were identified in the *Neurospora* genome, which could be classified into three groups on the basis of domain structure (Fig. 6). The first class contains genes similar to DHN-melanin PKS genes of the fungi *Exophila dermatitidis*⁶¹, *Colletotrichum lagenarium*⁶² and *Alternaria alternata*⁶³. Sequence identity to numerous expressed sequence tag (EST) sequences from sexual and perithecial libraries suggest a role in melanin pigment synthesis during sexual development⁶⁴. The genes in the second class are similar in structure to several fungal PKSs, including the *Aspergillus terreus lovF* gene required for lovastatin synthesis. The genes in the third class resemble other fungal genes, including the *A. terreus lovB* gene, which is also required for lovastatin synthesis.

Diterpene metabolism

Diterpenes comprise a diverse group of compounds, primarily in plants and fungi, with roles in defence, pathogenicity and regulation of plant growth. The genome sequence revealed several genes associated with diterpene biosynthesis in other organisms, including a terpene synthase, several genes related to gibberellin oxidases, and a member of the cytochrome P450 mono-oxygenase gene family. These genes include at least one member of each of the three enzyme classes required for the biosynthesis of gibberellic acid. Gibberellic acid, a normal growth regulator in plants, was first identified as a metabolic product of the plant pathogen *Gibberella fujikuroi*, a relative of *Neurospora* that causes 'foolish seedling' disease in rice⁶⁵. The presence of these genes in *Neurospora* suggests that many components necessary for gibberellic acid production were present in the ancestors of *Neurospora* and *G. fujikuroi*.

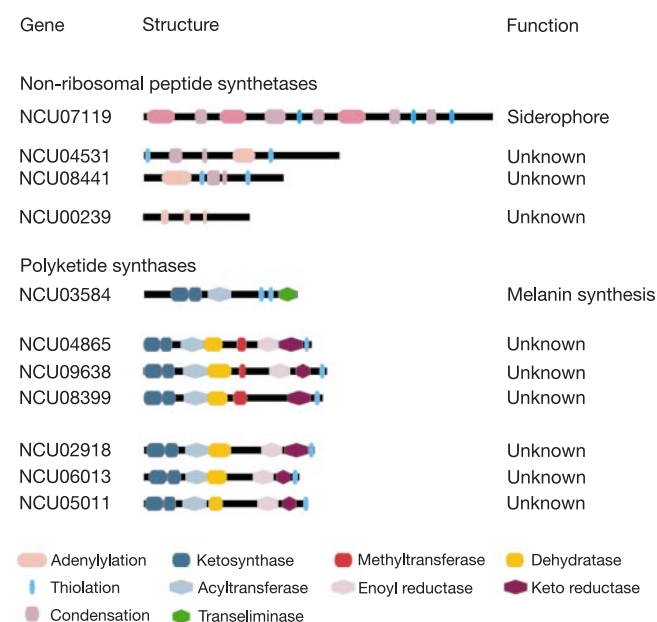


Figure 6 Domain structures of predicted *Neurospora* non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) genes. Domains were predicted using a combination of PFAM searches using HMMER, protein alignments and manual inspection.

We speculate that the secondary metabolism genes identified may have roles in morphogenesis and chemotropism⁶⁶, interspecies communication and possibly even chemical defence. The identification of these genes in *Neurospora* suggests that apparent major differences in lifestyles among related fungi, such as pathogenicity, may derive in part from minor modifications of gene function and expression.

Plant pathogenicity and *Neurospora*

The ability to parasitize living plants is widespread throughout the fungal kingdom. Although *Neurospora* is a saprotroph (that is, it feeds on dead or decaying matter), the genome sequence contains numerous genes similar to those required for plant pathogenesis identified in fungal pathogens. In particular, a number of genes were identified that have no known function in other organisms except in pathogenesis (Supplementary S8). *Neurospora* also possesses a wide range of extracellular enzymes capable of digesting plant cell wall polymers, although there is no clear cutinase homologue. Cutin is one of the main layers protecting the epidermis of the leaves of plants, and many, but not all, plant pathogens have cutinase activity. *Neurospora* has a wide range of cytochrome P450 enzymes that are important in some host–pathogen systems for detoxification of plant anti-fungal compounds. In addition, a large number of identified ABC (ATP-binding cassette) and MFS drug efflux systems could have a role in combating toxic plant compounds. The capability to form secondary metabolite members of the PKS, NRPS and terpenoid families, as described above, is present. Also, *Neurospora* contains all signal transduction components implicated in ascomycete pathogenesis that have been described so far. Thus, although *Neurospora* is not known to be a pathogen, the genome sequence has revealed many genes with similarity to those required for pathogenesis.

Discussion

Although *Neurospora* has been studied intensely for over 70 years, the analysis of the genome sequence has provided many new insights into a variety of cellular processes, including cell signalling, growth and differentiation, secondary metabolism and genome defence. The analysis has also uncovered surprising similarities between the saprotrophic *Neurospora* and pathogenic fungi, providing a new perspective on the molecular underpinnings of these lifestyles. Finally, the genome sequence has revealed the remarkable impact of RIP on the evolution of genes in *Neurospora*. Recent reports indicating the apparent presence of RIP in other fungi^{67,68} broaden the implications of our findings. The apparent lack of functional gene duplication in *Neurospora* provides a unique opportunity to study other modes of evolution in this experimentally tractable organism.

The genome sequence of *Neurospora* provides only a first glimpse into the genomic basis of the biological diversity of the filamentous fungi. Fungal genome sequences from the many ongoing⁶⁹ and planned⁷⁰ projects will expand this view as well as provide extraordinary opportunities for comparative analyses. This new era in fungal biology promises to yield insight into this important group of organisms, as well as to provide a deeper understanding of the fundamental cellular processes common to all eukaryotes. □

Methods

Strain and growth conditions

Twenty 5-ml cultures of *N. crassa* wild-type strain N150 (74-OR23-1VA; Fungal Genetics Stock Center 2489) were grown on a shaker in Vogel's minimal medium⁷ for 3 days at 32 °C. Tissues were collected, freeze-dried overnight and DNA was extracted as previously described⁷¹. DNA from the 20 samples was mixed and used for library construction.

Sequencing and assembly

The genome was sequenced by the WGS method. Plasmid (4-kb inserts) and fosmid (40-kb inserts) libraries were generated as described at <http://www-genome.wi.mit.edu/>. Jumping clone (subclone) libraries with 50-kb inserts were generated as described

elsewhere⁷². *Neurospora* cosmid and BAC clones were obtained from previously constructed libraries^{11,21}. Sequencing methods for all clone types are described at <http://www-genome.wi.mit.edu/>. All inserts were sequenced from both ends to generate paired reads. The sequence coverage generated is shown in Supplementary Information. The sequence was assembled using Arachne⁶. Finished sequence from linkage groups II and V was provided by MIPS and is available at <http://mips.gsf.de/proj/neurospora/>.

Annotation and analysis

We annotated the *Neurospora* genome using the Calhoun annotation system. The genome sequence was searched against the public protein databases using BLASTX with a threshold value of $E \leq 10^{-5}$. Genes were predicted using a combination of FGENESH, FGENESH+ and Genewise (Supplementary Information). The gene calling programs were validated against a test set of 191 previously characterized *Neurospora* proteins. Predicted genes were validated against ESTs aligned to the genome using SIM4. All predicted genes were searched against the PFAM set of hidden Markov models using the HMMER program and the public protein databases using BLASTP. Transfer RNAs were identified using the tRNAscan-SE program. Multigene families were constructed by searching each annotated gene against every other gene using BLASTP, requiring matches with $E \leq 10^{-5}$ over 60% of the longer gene length, and clustering genes based on single linkage transitive closure. Repeat sequences were detected by searching the genome sequence against itself using CrossMatch, filtering for alignments longer than 200 bp in length, and clustering pairs based on region overlap. Relics of RIP-mutated mobile elements were annotated by manual inspection.

The tree of MFS sugar transporters was created by aligning amino acid sequences using ClustalW, manually trimming to remove ambiguously aligned regions, and applying a neighbour-joining algorithm using PAUP*. RIP-mutated regions were detected by calculating one or both of two different dinucleotide ratios for sequence regions¹⁵. Regions with $\text{TpA}/\text{ApT} > 2$ or $(\text{CpA} + \text{TpG})/(\text{ApC} + \text{GpT}) < 0.7$ were predicted as RIP-mutated. Prediction of RIP sequence across the genome used only the TpA/ApT ratio, whereas the analysis of coding sequences used both (with a positive prediction by either measure taken as a prediction of RIP). RIP simulations were implemented in Matlab and were based on parameters derived from Table 2 of ref. 16. The simulation code is available on request. During each round of simulated RIP, every cytosine-containing dinucleotide on one strand (selected with equal probabilities) was mutated according to the probabilities: (CA = 0.3, CT = 0.05, CG = 0.01, CC = 0.009). DNA methylation was predicted by calculating the proportion of plasmid reads overlapping 1-kb windows from both the methylation-tolerant and methylation-intolerant libraries. Regions with greater than 70% of reads derived from the methylation-tolerant library were predicted to be methylated. Specificity was estimated as described in the text. Methylation was experimentally assessed using Southern analyses as described elsewhere³⁰. Sensitivity was estimated by testing 19 repetitive and RIP-mutated 1-kb regions that were not predicted to be methylated. Of the 19 regions, 14 were in fact methylated. Thus the data provide good specificity but poor sensitivity.

Predicted RNA-silencing genes were aligned with homologues from plants, animals and other fungi using T-Coffee v1.37. C-terminal and amino-terminal regions of low homology were removed and the sequences realigned until alignments started and stopped at regions of unambiguous similarity. Both neighbour-joining trees, using ClustalX, and maximum posterior probability trees, using MrBayes 2.01, were generated and analysed.

Analysis of predicted non-ribosomal peptide synthetases and polyketide synthases made use of genome data for *C. heterostrophus*, *Botryotinia fuckeliana*, *G. verticillioidea* and *G. zeae* provided by the Torrey Mesa Research Institute/Syngenta. Additional details, analysis results and the genome sequence are available at <http://www-genome.wi.mit.edu/>.

Received 24 December 2002; accepted 14 March 2003; doi:10.1038/nature01554.

- Payen, A. (rapporteur) Extrait d'un rapport adressé à M. Le Maréchal Duc de Dalmatie, Ministre de la Guerre, Président du Conseil, sur une altération extraordinaire du pain de munition. *Ann. Chim. Phys.* 3 Ser. 9, 5–21 (1843).
- Shear, C. L. & Dodge, B. O. Life histories and heterothallism of the red bread-mold fungi of the *Monilia strophila* group. *J. Agric. Res.* 34, 1019–1042 (1927).
- Lindgren, C. C. A six-point map of the sex chromosome of *Neurospora crassa*. *J. Genet.* 32, 243–256 (1936).
- Beadle, G. W. & Tatum, E. L. Genetic control of biochemical reactions in *Neurospora*. *Proc. Natl Acad. Sci. USA* 27, 499–506 (1941).
- Davis, R. H. *Neurospora: Contributions of a Model Organism* (Oxford Univ. Press, New York, 2000).
- Davis, R. H. & Perkins, D. D. *Neurospora: a model of model microbes*. *Nature Rev. Genet.* 3, 397–403 (2002).
- Perkins, D. D., Radford, A. & Sachs, M. S. *The Neurospora Compendium: Chromosomal Loci* (Academic, San Diego, California, 2001).
- Batzoglou, S. et al. ARACHNE: a whole-genome shotgun assembler. *Genome Res.* 12, 177–189 (2002).
- Schulte, U., Becker, I., Mewes, H. W. & Mannhaupt, G. Large scale analysis of sequences from *Neurospora crassa*. *J. Biotechnol.* 94, 3–13 (2002).
- Free, S. J., Rice, P. W. & Metzner, R. L. Arrangement of the genes coding for ribosomal ribonucleic acids in *Neurospora crassa*. *J. Bacteriol.* 137, 1219–1226 (1979).
- Aign, V., Schulte, U. & Hoheisel, J. D. Hybridization-based mapping of *Neurospora crassa* linkage groups II and V. *Genetics* 157, 1015–1020 (2001).
- Selker, E. U., Cambareri, E. B., Jensen, B. C. & Haack, K. R. Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* 51, 741–752 (1987).
- Selker, E. U. Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annu. Rev. Genet.* 24, 579–613 (1990).
- Cambareri, E. B., Jensen, B. C., Schabach, E. & Selker, E. U. Repeat-induced G-C to A-T mutations in *Neurospora*. *Science* 244, 1571–1575 (1989).
- Margolin, B. S. et al. A methylated *Neurospora* 5S rRNA pseudogene contains a transposable element inactivated by repeat-induced point mutation. *Genetics* 149, 1787–1797 (1998).

- Watters, M. K., Randall, T. A., Margolin, B. S., Selker, E. U. & Stadler, D. R. Action of repeat-induced point mutation on both strands of a duplex and on tandem duplications of various sizes in *Neurospora*. *Genetics* 153, 705–714 (1999).
- Cambareri, E. B., Singer, M. J. & Selker, E. U. Recurrence of repeat-induced point mutation (RIP) in *Neurospora crassa*. *Genetics* 127, 699–710 (1991).
- Rountree, M. R. & Selker, E. U. DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. *Genes Dev.* 11, 2383–2395 (1997).
- Irelan, J. T., Hagemann, A. T. & Selker, E. U. High frequency repeat-induced point mutation (RIP) is not associated with efficient recombination in *Neurospora*. *Genetics* 138, 1093–1103 (1994).
- Perkins, D. D., Margolin, B. S., Selker, E. U. & Haedo, S. D. Occurrence of repeat induced point mutation in long segmental duplications of *Neurospora*. *Genetics* 147, 125–136 (1997).
- Kelkar, H. S. et al. The *Neurospora crassa* genome: cosmid libraries sorted by chromosome. *Genetics* 157, 979–990 (2001).
- Hutchison, C. A. et al. Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* 286, 2165–2169 (1999).
- Ohno, S. *Evolution by Gene Duplication* (Springer, New York, 1970).
- Selker, E. U. et al. Dispersed 5S rRNA genes in *N. crassa*: structure, expression and evolution. *Cell* 24, 819–828 (1981).
- Selker, E. U. et al. The methylated component of the *Neurospora crassa* genome. *Nature* 422, 893–897 (2003).
- Kouzminova, E. & Selker, E. U. dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* 20, 4309–4323 (2001).
- Freitag, M., Williams, R. L., Kothe, G. O. & Selker, E. U. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* 99, 8802–8807 (2002).
- Russell, P. J., Rodland, K. D., Rachlin, E. M. & McCloskey, J. A. Differential DNA methylation during the vegetative life cycle of *Neurospora crassa*. *J. Bacteriol.* 169, 2902–2905 (1987).
- Foss, H. M., Roberts, C. J., Claeys, K. M. & Selker, E. U. Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation. *Science* 262, 1737–1741 (1993).
- Selker, E. U., Fritz, D. Y. & Singer, M. J. Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in *Neurospora*. *Science* 262, 1724–1728 (1993).
- Singer, M. J., Marcotte, B. A. & Selker, E. U. DNA methylation associated with repeat-induced point mutation in *Neurospora crassa*. *Mol. Cell. Biol.* 15, 5586–5597 (1995).
- Wolffe, A. P. & Matzke, M. A. Epigenetics: regulation through repression. *Science* 286, 481–486 (1999).
- Zamore, P. D. Ancient pathways programmed by small RNAs. *Science* 296, 1265–1269 (2002).
- Hutvagner, G. & Zamore, P. D. RNAi: nature abhors a double-strand. *Curr. Opin. Genet. Dev.* 12, 225–232 (2002).
- Catalanotto, C., Azzalin, G., Macino, G. & Cogoni, C. Involvement of small RNAs and role of the *qde* genes in the gene silencing pathway in *Neurospora*. *Genes Dev.* 16, 790–795 (2002).
- Aramayo, R. & Metzner, R. L. Meiotic transvection in fungi. *Cell* 86, 103–113 (1996).
- Shiu, P. K., Raju, N. B., Zickler, D. & Metzner, R. L. Meiotic silencing by unpaired DNA. *Cell* 107, 905–916 (2001).
- Shiu, P. K. & Metzner, R. L. Meiotic silencing by unpaired DNA. Properties, regulation and suppression. *Genetics* 161, 1483–1495 (2002).
- Volpe, T. A. et al. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833–1837 (2002).
- Lee, D., Pratt, R. J., McLaughlin, M., Aramayo, R. An Argonaute-like protein is required for meiotic silencing. *Genetics* (in the press).
- Linden, H., Ballario, P., Arpaia, G. & Macino, G. Seeing the light: news in *Neurospora* blue light signal transduction. *Adv. Genet.* 41, 35–54 (1999).
- Devlin, P. F. & Kay, S. A. Circadian photoperception. *Annu. Rev. Physiol.* 63, 677–694 (2001).
- Gustin, M. C., Albertyn, J., Alexander, M. & Davenport, K. MAP kinase pathways in the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 62, 1264–1300 (1998).
- Pott, G. B., Miller, T. K., Bartlett, J. A., Palas, J. S. & Selitrennikoff, C. P. The isolation of *FOS-1*, a gene encoding a putative two-component histidine kinase from *Aspergillus fumigatus*. *Fungal Genet. Biol.* 31, 55–67 (2000).
- Dohlman, H. G., Thorne, J., Caron, M. G. & Lefkowitz, R. J. Model systems for the study of seven-transmembrane-segment receptors. *Annu. Rev. Biochem.* 60, 653–688 (1991).
- Klein, P. S. et al. A chemoattractant receptor controls development in *Dictyostelium discoideum*. *Science* 241, 1467–1472 (1988).
- Josefsson, L. G. & Rask, L. Cloning of a putative G-protein-coupled receptor from *Arabidopsis thaliana*. *Eur. J. Biochem.* 249, 415–420 (1997).
- Aubry, L. & Firtel, R. Integration of signalling networks that regulate *Dictyostelium* differentiation. *Annu. Rev. Cell Dev. Biol.* 15, 469–517 (1999).
- Ivey, F. D., Yang, Q. & Borkovich, K. A. Positive regulation of adenylyl cyclase activity by a Go α homolog in *Neurospora crassa*. *Fungal Genet. Biol.* 26, 48–61 (1999).
- Gadd, G. M. in *The Growing Fungus* (eds Gow, N. A. R. & Gadd, G. M.) 183–210 (Chapman & Hall, London, 1994).
- Bootman, M. D. et al. Calcium signalling—an overview. *Semin. Cell Dev. Biol.* 12, 3–10 (2001).
- Lakin-Thomas, P. L. Effects of inositol starvation on the levels of inositol phosphates and inositol lipids in *Neurospora crassa*. *Biochem. J.* 292, 805–811 (1993).
- Silverman-Gavrila, L. B. & Lew, R. R. Regulation of the tip-high $[\text{Ca}^{2+}]$ gradient in growing hyphae of the fungus *Neurospora crassa*. *Eur. J. Cell Biol.* 80, 379–390 (2001).
- Cornelius, G., Gebauer, G. & Techel, D. Inositol trisphosphate induces calcium release from *Neurospora crassa* vacuoles. *Biochem. Biophys. Res. Commun.* 162, 852–856 (1989).
- Gimeno, C. J., Ljungdahl, P. O., Styles, C. A. & Fink, G. R. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* 68, 1077–1090 (1992).
- Lengeler, K. B. et al. Signal transduction cascades regulating fungal development and virulence. *Microbiol. Mol. Biol. Rev.* 64, 746–785 (2000).
- Adams, T. H., Wieser, J. K. & Yu, J. H. Asexual sporulation in *Aspergillus nidulans*. *Microbiol. Mol. Biol. Rev.* 62, 35–54 (1998).
- Ivey, F. D., Kays, A. M. & Borkovich, K. A. Shared and independent roles for a Go α protein and adenylyl cyclase in regulating development and stress responses in *Neurospora crassa*. *Eukaryot. Cell* 1, 634–642 (2002).

59. Yuan, W., Gentil, G., Budde, A. & Leong, S. Characterization of the *Ustilago maydis* *sid2* gene encoding a multidomain peptide synthetase in the ferrichrome biosynthetic gene cluster. *J. Bacteriol.* **183**, 4040–4051 (2001).
60. Lu, S.-W., Yoder, O. & Turgeon, B. G. Highly conserved virulence-related CPS1 homologs from plant and human pathogenic fungi. *Fungal Genet. News.* **48**, A246 (2001).
61. Feng, B. *et al.* Molecular cloning and characterization of *WdPKS1*, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in *Wangiella* (*Exophiala*) *dermatitidis*. *Infect. Immun.* **69**, 1781–1794 (2001).
62. Takano, Y. *et al.* Structural analysis of PKS1, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*. *Mol. Gen. Genet.* **249**, 162–167 (1995).
63. Takano, Y., Kubo, Y., Kawamura, C., Tsuge, T. & Furusawa, I. The *Alternaria alternata* melanin biosynthesis gene restores appressorial melanization and penetration of cellulose membranes in the melanin-deficient albino mutant of *Colletotrichum lagenarium*. *Fungal Genet. Biol.* **21**, 131–140 (1997).
64. Nelson, M. A. *et al.* Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*. *Fungal Genet. Biol.* **21**, 348–363 (1997).
65. Kurosawa, E. Experimental studies on the nature of the substance excreted by the 'bakanae' fungus. *Trans. Nat. Hist. Soc. Formosa* **16**, 213–227 (1926) (in Japanese).
66. Vigfusson, N. & Cano, R. Artificial induction of the sexual cycle of *Neurospora crassa*. *Nature* **249**, 383–385 (1974).
67. Graia, F. *et al.* Genome quality control: RIP (repeat-induced point mutation) comes to Podospora. *Mol. Microbiol.* **40**, 586–595 (2001).
68. Ikeda, K. *et al.* Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. *Mol. Microbiol.* **45**, 1355–1364 (2002).
69. TIGR Microbial Database (<http://www.tigr.org/tdb/mdb/mdbinprogress.html>) (2003).
70. Fungal Genome Initiative (<http://www-genome.wi.mit.edu/seq/fgi/>) (2002).
71. Luo, Z., Freitag, M. & Sachs, M. S. Translational regulation in response to changes in amino acid availability in *Neurospora crassa*. *Mol. Cell. Biol.* **15**, 5235–5245 (1995).
72. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
73. Pickford, A., Braccini, L., Macino, G. & Cogoni, C. The QDE-3 homologue RecQ-2 co-operates with QDE-3 in DNA repair in *Neurospora crassa*. *Curr. Genet.* **42**, 220–227 (2003).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements The authors would like to thank J. Arnold, H. Inoue, G. Turner, B. Bowman, P. Harriman and P. Youderian for their support. We also thank Lion Bioscience, V. Aign and J. Hoheisel for making available the BAC libraries used during sequencing; M. Karasz for developing the web pages hosting the *Neurospora* Interpro analysis at MIPS; the Torrey Mesa Research Institute/Syngenta for providing genome data for *C. heterostrophus*, *B. fuckeliana*, *G. verticillioidea* and *G. zeae*; and all members of the Whitehead Institute/MIT Center for Genome Research sequencing group. Funding for the *Neurospora* genome project was provided by the National Science Foundation. Additional funding was provided by the Deutsche Forschungsgemeinschaft, The Israel Science Foundation and the National Institutes of Health.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.E.G. (e-mail: jgalag@mit.edu) or B.B. (e-mail: bwb@genome.wi.mit.edu). The whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the project accession AABX00000000. The version described in this paper is the first version, AABX01000000.

Affiliations for authors: 1, Whitehead Institute Center for Genome Research, 320 Charles Street, Cambridge, Massachusetts 02141, USA; 2, Department of Plant Pathology, University of California, Riverside, California 92521, USA; 3, Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA; 4, Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JH, UK; 5, Celera Genomics, Rockville, Maryland 20850, USA; 6, Department of Biology, Texas A & M University, College Station, Texas 77843, USA; 7, Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA; 8, Department of Cellular and Structural Biology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA; 9, Department of Microbiology, University of Kansas Medical School, Kansas City, Kansas 66160, USA; 10, Department of Zoology, University of Florida, Gainesville, Florida 32611-8525, USA; 11, Department of Plant Pathology and Microbiology, Faculty of Agricultural, Food and Environmental Quality Sciences, The Hebrew University of Jerusalem, Rehovot 76100, Israel; 12, Institute of Biochemistry, Heinrich Heine University, 40225 Düsseldorf, Germany; 13, Laboratory of Plant Molecular Biology, Rockefeller University, New York, New York 10021, USA; 14, Department of Genome Oriented Bioinformatics, Technical University of Munich, Wissenschaftszentrum Weihenstephan, 85350 Freising-Weihenstephan, Germany; 15, Institute for Bioinformatics (MIPS), GSF-National Research Center for Environment and Health, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany; 16, T.H. Morgan School of Biological Sciences, University of Kentucky, Lexington, Kentucky 40506, USA; 17, The Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712, USA; 18, The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20878, USA; 19, Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, USA; 20, Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA; 21, Dipartimento Biotecnologie Cellulari ed Ematologia, Università di Roma La Sapienza, Rome, Italy; 22, School of Biological Sciences, Flinders University, P.O. Box 2100, Adelaide, South Australia 5001; 23, Department of Molecular Genetics, The Ohio State University, Columbus, Ohio 43210, USA; 24, Department of Molecular Genetics, The Ohio State University, Columbus, Ohio 43211, USA; 25, Department of Plant Pathology, University of Arizona, Tucson, Arizona 85721, USA; 26, School of Biological Sciences, University of Missouri-Kansas City, Kansas City, Missouri 64110, USA; 27, Department of Genetics, Dartmouth Medical School, Hanover, New Hampshire 03755, USA; 28, School of Biology, Leeds University, Leeds LS2 9JT, UK; 29, Department of Chemistry, California State Polytechnic University Pomona, Pomona, California 91768, USA; 30, Department of Plant Pathology and Microbiology, Texas A & M University, College Station, Texas 77843, USA; 31, Department of Environmental and Biomolecular Systems, OGI School of Science and Engineering, Oregon Health and Science University, Beaverton, Oregon 97006, USA; 32, Department of Biology, MIT, Cambridge, Massachusetts 02139, USA

Low-mass relics of early star formation

R. Schneider*†, A. Ferrara‡, R. Salvaterra‡, K. Omukai§ & V. Bromm||

* Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy

† “Enrico Fermi” Centre, Via Panisperna 89/A, 00184 Roma, Italy

‡ SISSA/International School for Advanced Studies, Via Beirut 4, 34100 Trieste, Italy

§ Division of Theoretical Astrophysics, National Astronomical Observatory, Mitaka, Tokyo 181-8588, Japan

|| Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

The earliest stars to form in the Universe were the first sources of light, heat and metals after the Big Bang. The products of their evolution will have had a profound impact on subsequent generations of stars. Recent studies^{1–7} of primordial star formation have shown that, in the absence of metals (elements heavier than helium), the formation of stars with masses 100 times that of the Sun would have been strongly favoured, and that low-mass stars could not have formed before a minimum level of metal enrichment had been reached. The value of this minimum level is very uncertain, but is likely to be between 10^{-6} and 10^{-4} that of the Sun^{6,8}. Here we show that the recent discovery⁹ of the most iron-poor star known indicates the presence of dust in extremely low-metallicity gas, and that this dust is crucial for the formation of lower-mass second-generation stars that could survive until today. The dust provides a pathway for cooling the gas that leads to fragmentation of the precursor molecular cloud into smaller clumps, which become the lower-mass stars.

The Hamburg/ESO objective prism survey team has recently reported the discovery⁹ of a star, HE0107–5240, with a mass of 0.8 solar masses, ($0.8 M_{\odot}$) and an iron abundance of $[\text{Fe}/\text{H}] = -5.3 \pm 0.2$ (here we adopt the standard definition $[A_i/A_j] = \log_{10}(N_i/N_j) - \log_{10}(N_i/N_j)_{\odot}$, where the subscript ‘ $\odot$ ’ refers to solar values). This represents the most iron-deficient star observed to date, and it has been found at a distance of about 11 kpc from the Sun—a region inaccessible to previous spectroscopic surveys that were limited to the inner halo of the Galaxy.

Does the existence of this $0.8 M_{\odot}$ very iron-deficient star suggest that the first generation of stars also contained low-mass and long-lived objects? The star HE0107–5240 does not show any sign of selective dust depletion that may have altered its elemental abundances. Being a low-mass star, elements heavier than Mg can not be produced by post-formation processes, and their chemical composition reflects the composition of the gas cloud out of which HE0107–5240 formed. The authors of ref. 9 thus conclude that the star is probably a second-generation object that formed from a gas cloud with metal abundance corresponding to $[\text{Fe}/\text{H}] = -5.3$ that has been pre-enriched by a supernova from a previous generation star.

As the metal composition of supernova ejecta depends on the initial mass of the progenitor star, the observed abundance pattern of HE0107–5240 may represent an important indication of the nature of the first stars and of the physical processes leading to low-mass star formation in primordial gas clouds.

Two plausible models need to be considered. If we assume pre-formation origin for all the heavy elements observed in HE0107–5240, the gas cloud out of which the star formed would have had to be pre-enriched to a metallicity as high as $Z = 10^{-2} Z_{\odot}$ (where $Z_{\odot}$ is the solar metallicity). This metallicity is significantly higher than the proposed minimum level of pre-enrichment $Z_{\text{cr}} = 10^{-5 \pm 1} Z_{\odot}$ (ref. 6) above which low-mass stars can form.

In the second model, suggested in ref. 9, the observed abundances

of C, N, Na and Mg may be due to post-formation mechanisms, such as mass transfer from a companion star, or self-enrichment according to recent nucleosynthesis models of stars of mass similar to that inferred for HE0107–5240 ($\sim 0.8 M_{\odot}$) of zero or near-zero initial metallicity¹⁰. In this “minimal pre-enrichment model”, only elements heavier than Mg have a pre-formation origin. Indeed, the authors of ref. 9 argue that the abundance pattern of elements heavier than Mg is consistent with the predicted elemental yields of a 25–25 $M_{\odot}$ star that exploded as a type II supernova, indicating that HE0107–5240 may have formed from a gas cloud which had been pre-enriched by such a supernova. However, if the first stars had a mass within the range $200 M_{\odot} \leq M \leq 220 M_{\odot}$ and exploded as pair-creation supernovae¹¹, the observed abundance pattern of HE0107–5240 for elements heavier than Mg can be equally well reproduced.

Thus, with the current data it is not possible to uniquely determine the mass of the first-generation star(s) that produced the metals locked up in HE0107–5240. To do this, we need improved upper limits on the abundances of Zn (which is predicted to be about two orders of magnitudes smaller in the ejecta of a 200–220 $M_{\odot}$ pair-creation supernova than for a 22 $M_{\odot}$ type II supernova) and heavy r-process elements in HE0107–5240, as well as in other extremely metal-poor stars to be found in the future. The r process is believed to only operate in an extremely neutron-rich environment¹², as is provided during the creation of a central neutron star in a core-collapse (type II) supernova. A pair-creation supernova, on the other hand, is fundamentally different¹¹, in that no remnant is created, and that the necessary condition for the r process is not realized. Observing the absence or presence of r-process elements, therefore, holds the promise of directly probing the nature of the first stars¹³.

Thus we can ask: what was the actual level of metallicity of the parent cloud out of which HE0107–5240 was born? Assuming the same interstellar medium dilution factor for all the observed heavy elements, we can directly derive a lower limit for the metallicity of the cloud. According to such estimates, the explosions of stars with mass in the ranges $22 M_{\odot}$ and $200 M_{\odot} < M_{\star} < 220 M_{\odot}$ were able to enrich the star-forming gas up to a metallicity $10^{-5.5} \leq Z/Z_{\odot} \leq 10^{-5.1}$. This range of initial metallicities falls within the proposed minimum level of metal pre-enrichment⁶ above which low-mass objects can form. Thus, the observation of HE0107–5240 appears to be extremely helpful in determining the critical conditions that finally enable low-mass star formation at early times. Following the analysis of ref. 6, we recall that if some efficient cooling mechanism is present, the energy deposited by gravitational contraction cannot balance the radiative losses, and the gas cloud cools and fragments into increasingly smaller gas clumps. The necessary conditions to stop fragmentation and start gravitational contraction within each clump are that cooling becomes inefficient, and the Jeans mass of the fragments does not decrease any further, thus favouring fragmentation into sub-clumps. This condition depends somewhat on the geometry of the fragments, and translates into $\gamma > 4/3$ ($\gamma > 1$) for spherical (filamentary) clumps, where γ is the adiabatic index defined as $T \propto n^{\gamma-1}$.

For a metal-free gas, the only efficient coolant is molecular hydrogen. Cooling due to molecular line emission becomes inefficient at densities above $n > 10^3 \text{ cm}^{-3}$, and fragmentation stops when the minimum fragment mass is of the order of 10^3 – $10^4 M_{\odot}$. As the gas becomes more and more enriched with heavy elements, the cooling rate at low density $n \leq 10^3 \text{ cm}^{-3}$ increases because of metal (especially O and C) line emission. More importantly, however, the depletion factor, defined as the dust-to-metals mass ratio $f_{\text{dep}} = M_{\text{gr}}/M_Z$, is responsible for activating (through dust-gas thermal exchanges) a new strong phase of fragmentation at high density, which finally leads to fragments of low mass. For initial metallicities in the range 10^{-4} – $1 Z_{\odot}$, the presence of dust becomes almost irrelevant, and the minimum fragment masses are in the range 10^{-2} – $1 M_{\odot}$, several orders of magnitude smaller than in the metal-free case.

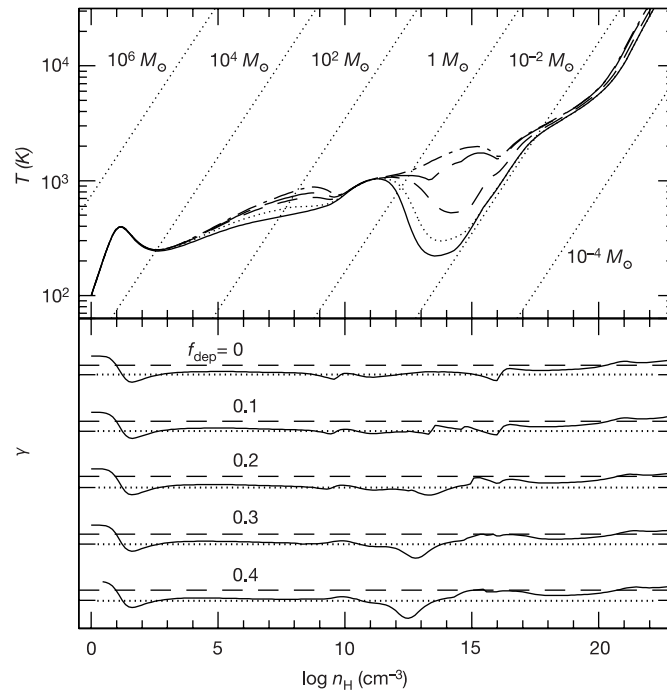


Figure 1 Collapse and fragmentation of low-metallicity star-forming gas clouds. Upper panel, curves show the temperature evolution as a function of the hydrogen number density of protostellar clouds with the same initial metallicity $Z = 10^{-5.1} Z_{\odot}$, consistent with the observed properties of HE0107–5240, but varying dust depletion factors: $f_{\text{dep}} = 0$ (dash-dotted), $f_{\text{dep}} = 0.1$ (long-dashed), $f_{\text{dep}} = 0.2$ (short-dashed),

$f_{\text{dep}} = 0.5$ (dotted) and $f_{\text{dep}} = 1$ (solid). The diagonal dotted lines correspond to constant Jeans mass for spherical clumps. Lower panel, the adiabatic index γ as a function of the hydrogen number density for the curves shown in the upper panel. Dotted (dashed) lines correspond to $\gamma = 1$ ($\gamma = 4/3$) (see text). Curves are labelled with the appropriate value of f_{dep} .

In Fig. 1, we show the evolution of a star-forming gas cloud with initial metallicity of $Z = 10^{-5.1} Z_{\odot}$ as if it were pre-enriched by the pair-creation supernova explosion of a $205 M_{\odot}$ star. The various curves correspond to different values for the depletion factor. The thermal evolution of the gas is insensitive to the presence of different amounts of dust grains as long as the density is below 10^5 cm^{-3} . In this regime, the cooling rate is dominated by molecular line cooling. At higher densities, gas clouds with depletion factors ≥ 0.2 experience a new phase of fragmentation (shown as the dip in the γ evolution at $n \approx 10^{12} \text{ cm}^{-3}$), which is due to dust–gas thermal exchanges and lasts for about three orders of magnitude in density, leading to characteristic fragment masses of the order of 10^{-2} – $10^{-1} M_{\odot}$.

Thus it appears that a gas cloud with initial metallicity of $Z \approx 10^{-5.1} Z_{\odot}$ and $f_{\text{dep}} \geq 0.2$ (for comparison the Galactic value is $f_{\text{dep}} = 0.47$) can generate low-mass fragments, subsequently leading to the formation of low-mass stars, consistent with the observation of HE0107–5240. As an alternative model, a mechanism has been proposed¹⁴ for the formation of a population of extremely metal-poor intermediate-mass stars at very high redshifts. Within this model, the formation of these stars is directly linked to the death of a very massive star exploding as a pair-creation supernova. Indeed, shock compression, heating and subsequent cooling to high density reduces the fragment mass in primordial gas to $\sim 10 M_{\odot}$, allowing small-mass stars to form. It might be that HE0107–5240 is the first observed member of such a population.

Our analysis shows that the discovery of HE0107–5240 does not conflict with recent studies which suggest that the formation of low-mass stars that can survive to the present day is inhibited before a minimum level of metal enrichment is reached. The actual value of this threshold metallicity within the currently uncertain range $Z_{\text{cr}} = 10^{-5.5 \pm 1} Z_{\odot}$ (ref. 6) strongly depends on the efficiency of dust formation in first-generation supernova ejecta.

If the previous lack of very iron-deficient halo stars is to be

ascribed to the brighter magnitude limits of previous surveys, it is likely that the Hamburg/ESO prism survey will lead to the identification of more stars with $[\text{Fe}/\text{H}] < -5$. The statistics and properties of these stars should provide insights into the first epochs of cosmic star formation. In particular, if observers were to find a large number of these objects, this might imply that second-generation stars were already forming with characteristic masses of $1 M_{\odot}$, as in the local Universe. Conversely, if future data were to show that the bulk of metal-poor halo stars have a metallicity in the range $Z = 10^{-4}$ – $10^{-3} Z_{\odot}$ with only a small number of $< 10^{-5} Z_{\odot}$ outliers, then this would indicate that in most cases small-mass stars were able to form only after this level of metallicity had been achieved. $\square$

Received 23 December 2002; accepted 10 March 2003; doi:10.1038/nature01579.

- Omukai, K. & Nishi, R. Formation of primordial protostars. *Astrophys. J.* **508**, 141–150 (1998).
- Abel, T., Bryan, G. & Norman, M. The formation of the first star in the universe. *Science* **295**, 93–98 (2002).
- Bromm, V., Coppi, P. S. & Larson, R. B. The formation of the first stars. I. The primordial star-forming cloud. *Astrophys. J.* **564**, 23–51 (2002).
- Ripamonti, E., Haardt, F., Ferrara, A. & Colpi, M. Radiation from the first forming stars. *Mon. Not. R. Astron. Soc.* **334**, 401–418 (2002).
- Nakamura, F. & Umemura, M. The stellar initial mass function in primordial galaxies. *Astrophys. J.* **569**, 549–557 (2002).
- Schneider, R., Ferrara, A., Natarajan, P. & Omukai, K. First stars, very massive black holes and metals. *Astrophys. J.* **579**, 30–39 (2002).
- Omukai, K. & Palla, F. Formation of the first stars by accretion. *Astrophys. J.* (submitted).
- Bromm, V., Ferrara, A., Coppi, P. S. & Larson, R. B. The fragmentation of pre-enriched primordial objects. *Mon. Not. R. Astron. Soc.* **328**, 969–976 (2001).
- Christlieb, N. et al. A stellar relic from the early Milky Way. *Nature* **419**, 904–906 (2002).
- Siess, L., Livio, M. & Lattanzio, J. Structure, evolution, and nucleosynthesis of primordial stars. *Astrophys. J.* **570**, 329–343 (2002).
- Heger, A. & Woosley, S. E. The nucleosynthesis signature of population III. *Astrophys. J.* **567**, 532–543 (2002).
- Truran, J. W., Cowan, J. J., Pilachowski, C. A. & Sneden, C. Probing the neutron-capture nucleosynthesis history of galactic matter. *Publ. Astron. Soc. Pacif.* **114**, 1293–1308 (2002).
- Qian, Y.-Z. & Wasserburg, G. J. Determination of nucleosynthetic yields of supernovae and very massive stars from abundances in metal-poor stars. *Astrophys. J.* **567**, 515–531 (2002).
- Mackey, J., Bromm, V. & Hernquist, L. Three epochs of star formation in the high redshift universe. *Astrophys. J.* **586**, 1–11 (2003).

Acknowledgements We acknowledge partial support from the Research and Training Network 'The Physics of the Intergalactic Medium' set up by the European Community.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.S. (e-mail: raffa@arcetri.astro.it).

First-generation black-hole-forming supernovae and the metal abundance pattern of a very iron-poor star

Hideyuki Umeda & Ken'ichi Nomoto

Department of Astronomy, School of Science, University of Tokyo, Tokyo 113-0033, Japan

It has been proposed¹ theoretically that the first generation of stars in the Universe (population III) would be as massive as 100 solar masses ($100 M_{\odot}$), because of inefficient cooling^{2–4} of the precursor gas clouds. Recently, the most iron-deficient (but still carbon-rich) low-mass star—HE0107–5240—was discovered⁵. If this is a population III star that gained its metals (elements heavier than helium) after its formation, it would challenge the theoretical picture of the formation of the first stars. Here we report that the patterns of elemental abundance in HE0107–5240 (and other extremely metal-poor stars) are in good accord with the nucleosynthesis that occurs in stars with masses of 20–130 $M_{\odot}$ when they become supernovae if, during the explosions, the ejecta undergo substantial mixing and fall-back to form massive black holes. Such supernovae have been observed⁷. The abundance patterns are not, however, consistent with enrichment by supernovae from stars in the range 130–300 $M_{\odot}$. We accordingly infer that the first-generation supernovae came mostly from explosions of ~ 20 –130 $M_{\odot}$ stars; some of these produced iron-poor but carbon- and oxygen-rich ejecta. Low-mass second-generation stars, like HE0107–5240, could

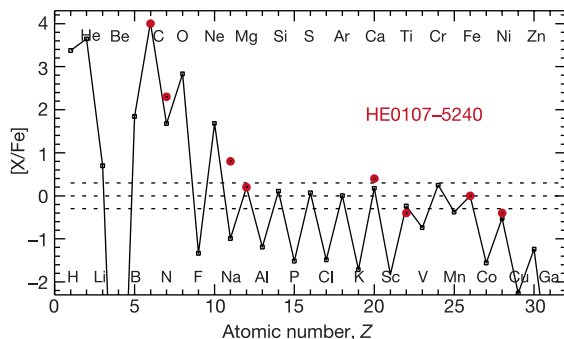


Figure 1 Elemental abundances of HE0107–5240, compared with a theoretical supernova yield. HE0107–5240 (filled circles) is the most Fe-deficient, C-rich star yet observed, with $[Fe/H] = -5.3$ and very large ratios of $[C/Fe] = 4.0$ and $[N/Fe] = 2.3$. Here the supernova model is the population III 25 $M_{\odot}$ core collapse, with relatively small explosion energy $E_{51} = E_{\text{exp}}/10^{51} \text{ erg} = 0.3$ (see Fig. 2). In this model, the supernova ejecta are assumed to be mixed inside the He core (within the mass coordinate $M_r = 1.83 M_{\odot}$ and $6.00 M_{\odot}$) and only a small fraction of the matter, 0.002% ($f = 0.00002$), is ejected from this region. The ejected Fe (or ^{56}Ni) mass, $8 \times 10^{-6} M_{\odot}$, is so small that the large C/Fe ratio can be realized.

form because the carbon and oxygen provided pathways for the gas to cool.

The abundance pattern of HE0107–5240 is characterized by very large C/Fe and N/Fe ratios, while the abundances of elements heavier than Mg are as low as that of Fe (Fig. 1). How could this peculiar abundance pattern have been formed? We note that this is not the only extremely metal-poor (EMP) star to have large C/Fe and N/Fe ratios—several other such stars have been discovered^{8–10}. Therefore a reasonable explanation of the abundance pattern should also apply to other EMP stars.

Before describing our model, we examine two alternative explanations of the abundance pattern of C-rich EMP stars⁵. These explanations assume that small-mass stars formed from EMP gases, and that carbon was enhanced after the formation. The first is mass transfer from a companion star. This explanation suffers from two difficulties. (1) Not all EMP stars have observational evidence for the existence of companion stars, though further observations to obtain radial velocity data are necessary¹¹. (2) Some EMP stars (for example, CS22949–037 and CS29498–043) have a relatively large Si abundance, $[Si/Fe] \approx 1$ (refs 8–10). (Here $[A/X] = \log_{10}(N_A/N_X) - \log_{10}(N_A/N_X)_{\odot}$, where the subscript $\odot$ refers to the Sun and N_A and N_X are the number densities of elements A and X respectively.) The overabundance of Si is very difficult to explain by mass transfer from companion AGB stars.

The second explanation is matter accretion onto a population III star during repeated passages through the Galactic disk⁶. This model alone cannot explain the large abundance of C and N—for example, in HE0107–5240, with $[C/H] = -1.3$. Therefore, C and N need to be produced in the interior of the EMP stars. Recently, nucleosynthesis in the low-mass population III stars was studied extensively^{12,13}. Such studies reveal that C and N abundances can be significantly enhanced if mixing is efficient during the He-flash in the core at the tip of the red-giant branch. However, HE0107–5240 is identified as a low-mass ($\sim 0.8 M_{\odot}$) red giant⁵, and thus this mechanism is not likely to operate.

Here we consider a model in which C-rich EMP stars are produced in the ejecta of (almost) metal-free supernovae mixed with EMP interstellar matter. We use population III pre-supernova progenitor models¹⁴, simulate the supernova explosion, and calcu-

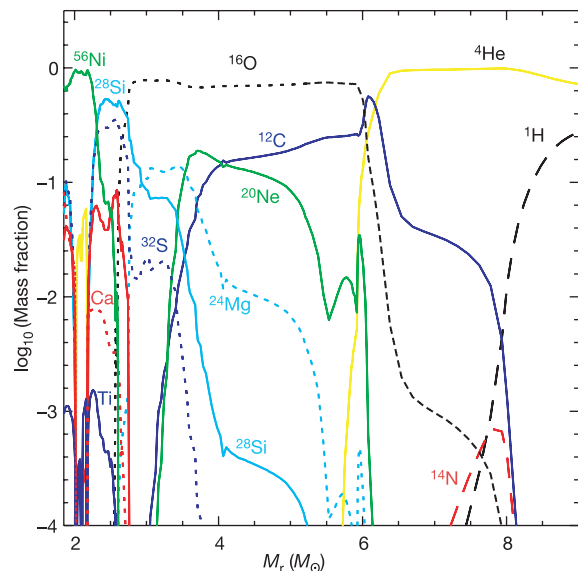


Figure 2 The post-explosion abundance distributions for the population III 25 $M_{\odot}$ model with explosion energy $E_{51} = 0.3$. Explosive nucleosynthesis takes place behind the shock wave that is generated at $M_r = 1.8 M_{\odot}$ and propagates outwards. The hydrogen-rich envelope at $M_r > 9 M_{\odot}$ is not shown.

late detailed nucleosynthesis. The elemental abundances of one of our models are in good agreement with HE0107–5240 (Fig. 1). In this model, the progenitor mass is $25 M_{\odot}$ and the explosion energy $E_{\text{exp}} = 3 \times 10^{50}$ erg (or $E_{51} = E_{\text{exp}}/10^{51} = 0.3$). The resultant abundance distribution is shown in Fig. 2. The processed material is assumed to mix uniformly in the region from $M_r = 1.8 M_{\odot}$ to $6.0 M_{\odot}$ (where M_r denotes the lagrangian mass coordinate measured from the centre of the pre-supernova model). Such a large-scale mixing was found to take place in SN1987A and various explosion models^{15,16}. Almost all material below $M_r = 6.0 M_{\odot}$ falls back to the central remnant and only a small fraction (factor $f = 2 \times 10^{-5}$) is ejected from this region.

The CNO elements in the ejecta were produced by pre-collapse He shell burning in the He layer, which contains $0.2 M_{\odot}^{12}\text{C}$. Mixing of H into the He shell-burning region produces $4 \times 10^{-4} M_{\odot}^{14}\text{N}$. On the other hand, only a small amount of heavier elements (Mg, Ca and Fe-peak elements) are ejected, and their abundance ratios are the average over the region of $M_r = 1.8$ – $6.0 M_{\odot}$. The subsolar ratios of $[\text{Ti}/\text{Fe}] = -0.4$ and $[\text{Ni}/\text{Fe}] = -0.4$ are the results of the relatively small explosion energy ($E_{51} = 0.3$). With this ‘mixing and fallback’, the large C/Fe and C/Mg ratios observed in HE0107–5240 are well reproduced. In this model, N/Fe appears to be underproduced. However, N can be produced inside the EMP stars through the C–N cycle, and brought up to the surface during the first dredge-up stage while becoming a red-giant star¹⁷.

This ‘mixing and fallback’ is commonly required to reproduce the abundance pattern of typical EMP stars¹⁴. In Fig. 3 we show a model that is in good agreement with CS22949 – 037. This star has $[\text{Fe}/\text{H}] \approx -4.0$ and is also C and N rich¹⁰, though C/Fe and N/Fe ratios are smaller than HE0107–5240. This model is the explosion of a $30 M_{\odot}$ star with $E_{51} = 20$. In this model, the mixing region ($M_r = 2.33$ – $8.56 M_{\odot}$) is chosen to be smaller than the entire He core ($M_r = 13.1 M_{\odot}$), in order to reproduce relatively large Mg/Fe and Si/Fe ratios. Similar degrees of mixing would also reproduce the abundances of CS29498 – 043 (ref. 9), which shows a similar abundance pattern. We assume a larger fraction of ejection, 2%, from the mixed region for CS22949 – 037 than HE0107–5240, because the C/Fe and N/Fe ratios are smaller. The ejected Fe mass is $0.003 M_{\odot}$. The larger explosion energy model is favoured for explaining the large Zn/Fe, Co/Fe and Ti/Fe ratios¹⁴. Without mixing, elements produced in the deep explosive burning regions, such as Zn, Co and Ti, would be underproduced. Without fallback, the abundance ratios of heavy elements to lighter elements, such as Fe/C, Fe/O and Mg/C would be too large.

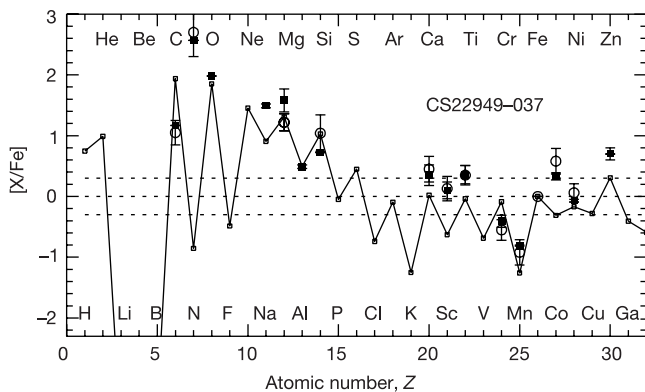


Figure 3 Elemental abundances of CS22949 – 037, compared with a theoretical supernova yield. The open circles and filled squares are observed data from refs 8 and 10, respectively. For the theoretical model (small open squares and solid lines), a high-energy model ($E_{51} = 20$) is adopted for the relatively large abundances of Co and Zn. In this $30 M_{\odot}$ model, the He core is $13.1 M_{\odot}$. The mixing region ($M_r = 2.33$ – $8.56 M_{\odot}$) is smaller than the entire He core, and the matter ejection factor $f = 0.002$ is larger than the model for HE0107–5240 because of larger Mg/C and C/Fe ratios.

In this model, Ti, Co, Zn and Sc are still under-produced, but these elements may be enhanced efficiently in aspherical explosions¹⁸. Almost the same effects as the ‘mixing and fallback mechanism’ are realized if the explosion is jet-like, although the total energy can be smaller if the beaming angle of the jet is small. Similarly, the ‘mixing and fallback’ process can reproduce the abundance pattern of the typical EMP stars without enhancement of C and N. For example, the averaged abundances of $[\text{Fe}/\text{H}] \approx -3.7$ stars in ref. 8 can be fitted well with the model of $25 M_{\odot}$ and $E_{51} = 20$ but with a larger fraction ($\sim 10\%$) of the processed materials in the ejecta.

In our model, $[\text{Fe}/\text{H}]$ of several kinds of EMP stars can be understood in the supernova-induced star formation scheme^{19,20}. In this scheme, $[\text{Fe}/\text{H}]$ of the second-generation stars are determined by the ejected Fe mass divided by the mass of hydrogen swept up by the supernova ejecta. As the swept-up hydrogen mass is roughly proportional to the explosion energy, $\text{Fe}/\text{H} \propto (M(\text{Fe})/0.07 M_{\odot})/E_{51}$, where $M(\text{Fe})$ is the ejected Fe mass. The average stars of $[\text{Fe}/\text{H}] \approx -3.7$ (ref. 7), CS22949 – 037 and HE0107–5240 correspond to $(M(\text{Fe})/0.07 M_{\odot})/E_{51} = 0.07$, 0.002 and 0.0004, respectively. This correspondence suggests that $[\text{Fe}/\text{H}]$ of the EMP stars does not reflect the age of the stars, but rather the properties of the supernovae, such as the degree of mixing and fallback or collimation of a jet.

We have shown that the ejecta of core-collapse supernova explosions of 20 – $130 M_{\odot}$ stars can well account for the abundance pattern of EMP stars. In contrast, the observed abundance patterns cannot be explained by the explosions of more massive, 130 – $300 M_{\odot}$ stars. These stars undergo pair-instability supernova and are disrupted completely^{14,21}, which cannot be consistent with the large C/Fe ratio observed in HE0107–5240 and other C-rich EMP stars. Also, the abundance ratios of iron-peak elements ($[\text{Zn}/\text{Fe}] < -0.8$ and $[\text{Co}/\text{Fe}] < -0.2$) in the ejecta of pair-instability supernovae^{14,21} cannot explain the large Zn/Fe and Co/Fe (ref. 14) in the typical EMP stars^{8,22,23} and CS22949 – 037. Therefore the supernova progenitors that are responsible for the formation of EMP stars are most likely in the range $M \approx 20$ – $130 M_{\odot}$, but not more massive than $130 M_{\odot}$.

We have also shown that the most iron-poor star, as well as other C-rich EMP stars, is likely to be enriched by massive supernovae that are characterized by relatively small Fe ejection. Such supernovae are not hypothetical, but have been observed, forming a distinct class of type II supernovae (‘faint supernovae’)²⁴. The prototype is SN1997D, which is very underluminous and shows quite narrow spectral lines²⁵: these features are well modelled as an explosion of a $25 M_{\odot}$ star ejecting only $2 \times 10^{-3} M_{\odot}^{56}\text{Ni}$ with small explosion energy $E_{51} \approx 0.4$ (ref. 7). SN1999br is a very similar faint supernova²⁵. On the other hand, typical EMP stars without enhancement of C and N correspond to the abundance pattern of energetic supernovae (‘hypernovae’²⁴). For both cases, black holes more massive than ~ 3 – $10 M_{\odot}$ must be left as results of fallback, suggesting copious formation of the first black holes from the first stars. These black holes may make up some of the dark mass in the Galactic halo.

In our model, HE0107–5240 with $[\text{Fe}/\text{H}] = -5.3$ was formed from C- and O-enhanced gases with $[\text{C}/\text{O}/\text{H}] \approx -1$. With such enhanced C and O, the cooling efficiency is large enough to form small-mass stars. In other words, our model predicts that low-mass EMP stars with $[\text{Fe}/\text{H}] < -4$ are likely to have enhanced $[\text{C}, \text{N}, \text{O}/\text{Fe}]$ and $[\text{Mg}/\text{Fe}]$ in some cases. A consequence of the low-mass EMP stars being carbon-rich is that the population III stars that provided their metals are massive enough to form (the first) black holes. □

Received 18 December 2002; accepted 4 March 2003; doi:10.1038/nature01571.

1. Abel, T., Bryan, G. L. & Norman, M. L. The formation of the first star in the Universe. *Science* **295**, 93–98 (2002).
2. Omukai, K. & Nishi, R. Formation of primordial protostars. *Astrophys. J.* **508**, 141–150 (1998).
3. Nakamura, F. & Umemura, M. On the mass of population III stars. *Astrophys. J.* **515**, 239–248 (1999).

4. Schneider, R., Ferrara, A., Natarajan, P. & Omukai, K. First stars, very massive black holes, and metals. *Astrophys. J.* **571**, 30–39 (2002).
5. Christlieb, N. *et al.* A stellar relic from the early Milky Way. *Nature* **419**, 904–906 (2002).
6. Yoshii, Y. Metal enrichment in the atmospheres of extremely metal-deficient dwarf stars by accretion of interstellar matter. *Astron. Astrophys.* **97**, 280–290 (1981).
7. Turatto, M. *et al.* The peculiar type II supernova 1997D. A case for a very low ^{56}Ni mass. *Astrophys. J.* **498**, L129–L133 (1998).
8. Norris, J. E., Ryan, S. G. & Beers, T. C. Extremely metal-poor stars. VIII. High-resolution, high signal-to-noise ratio analysis of five stars with $[\text{Fe}/\text{H}] < -3.5$. *Astrophys. J.* **561**, 1034–1059 (2001).
9. Aoki, W., Ryan, S. G., Beers, T. C. & Ando, H. The chemical composition of carbon-rich, very metal poor stars. *Astrophys. J.* **567**, 1166–1182 (2002).
10. Depagne, E. *et al.* First stars II. Elemental abundances in the extremely metal-poor star CS 22949-037. *Astron. Astrophys.* **390**, 187–198 (2002).
11. Ryan, S. G. Carbon-rich, extremely metal-poor population III stars. In *CNO in the Universe* (eds Charbonnel, C., Schaerer, D. & Meynet, G.) (in the press); preprint astro-ph/0211608 at (<http://xxx.lanl.gov>) (2002).
12. Fujimoto, M. Y., Ikeda, Y. & Iben, I. Jr The origin of extremely metal-poor stars and the search for population III. *Astrophys. J.* **529**, L25–L28 (2000).
13. Siess, L., Livio, M. & Lattanzio, J. Structure, evolution, and nucleosynthesis of primordial stars. *Astrophys. J.* **570**, 329–343 (2002).
14. Umeda, H. & Nomoto, K. Nucleosynthesis of zinc and iron peak elements in population III type II supernovae. *Astrophys. J.* **565**, 385–404 (2002).
15. Hachisu, I., Matsuda, T., Nomoto, K. & Shigeyama, T. Non linear growth of Rayleigh–Taylor instabilities and mixing. *Astrophys. J.* **358**, L57–L61 (1990).
16. Kifonidis, K., Plewa, T., Janka, H.-Th. & Müller, E. Nucleosynthesis and clump formation in a core-collapse supernova. *Astrophys. J.* **531**, L123–L126 (2000).
17. Boothroyd, A. I. & Sackmann, I.-J. The CNO isotopes: Deep circulation in red giants and first and second dredge-up. *Astrophys. J.* **510**, 217–231 (1999).
18. Maeda, K. & Nomoto, K. Bipolar supernova explosions: nucleosynthesis and implication on abundances in extremely metal-poor stars. *Astrophys. J.* (submitted); preprint available at (<http://xxx.lanl.gov/astro-ph/0304172>) (2003).
19. Audouze, J. & Silk, J. The first generation of stars: first steps toward chemical evolution of galaxies. *Astrophys. J.* **451**, L49–L52 (1995).
20. Shigeyama, T. & Tsujimoto, T. Fossil imprints of the first-generation supernova ejecta in extremely metal-deficient stars. *Astrophys. J.* **507**, L135–L139 (1998).
21. Heger, A. & Woosley, S. E. The nucleosynthetic signature of population III. *Astrophys. J.* **567**, 532–543 (2002).
22. McWilliam, A., Preston, G. W., Sneden, C. & Searle, L. Spectroscopic analysis of 33 of the most metal poor stars II. *Astron. J.* **109**, 2757–2799 (1995).
23. Blake, L. A. J., Ryan, S. G., Norris, J. E. & Beers, T. C. Neutron-capture elements in the Sr-rich, Ba-normal metal-poor giant CS22897-008. *Nucl. Phys. A* **688**, 502–504 (2001).
24. Nomoto, K., *et al.* Hypernovae and their nucleosynthesis. In *A Massive Star Odyssey, from Main Sequence to Supernova* (eds van der Hucht, K. A., Herrero, A. & Esteban, C.) (in the press); preprint astro-ph/0209064 at (<http://xxx.lanl.gov>) (2002).
25. Zampieri, L. *et al.* Peculiar, low luminosity type II supernovae: Low energy explosions in massive progenitors? *Mon. Not. R. Astron. Soc.* **338**, 711–716 (2003).

Acknowledgements This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science, Culture, Sports and Technology in Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.N. (e-mail: nomoto@astron.s.u-tokyo.ac.jp).

No mixing of superconductivity and antiferromagnetism in a high-temperature superconductor

I. Bozovic^{*†}, G. Logvenov^{*}, M. A. J. Verhoeven^{*}, P. Caputo^{*}, E. Goldobin^{*} & T. H. Geballe[‡]

^{*} Oxxel GmbH, Technologiepark Universitaet, Bremen 28359, Germany

[†] Geballe Laboratory for Advanced Materials, [‡] Applied Physics Department, Stanford University, Stanford, California 94305, USA

There is still no universally accepted theory of high-temperature superconductivity. Most models assume that doping creates ‘holes’ in the valence band of an insulating, antiferromagnetic ‘parent’ compound, and that antiferromagnetism and high-temperature superconductivity are intimately related^{1–8}. If their respective energies are nearly equal, strong antiferromagnetic

fluctuations (temporally and spatially restricted antiferromagnetic domains) would be expected in the superconductive phase, and superconducting fluctuations would be expected in the antiferromagnetic phase⁷; the two states should ‘mix’ over an extended length scale⁸. Here we report that one-unit-cell-thick antiferromagnetic La_2CuO_4 barrier layers remain highly insulating and completely block a supercurrent; the characteristic decay length is 1 Å, indicating that the two phases do not mix. We likewise found that isolated one-unit-cell-thick layers of $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ remain superconducting. The latter further implies that, on doping, new electronic states are created near the middle of the bandgap. These two findings are in conflict with most proposed models, with a few notable exceptions that include postulated spin–charge separation².

In a radical departure from the conventional theory of metals and superconductivity, Anderson proposed^{1,2} that the basic physics of high-temperature superconductivity (HTS) is captured in a single-band Hubbard model hamiltonian. This approach has pervaded theoretical thinking about copper oxides. For zero (or very low) doping, the ground state of this hamiltonian should be an antiferromagnetic (AF) Mott–Hubbard insulator (MHI); for the chemical potential μ larger than some critical value μ_c , HTS should occur. However, the model has not been solved analytically, and specific predictions are frequently made at the expense of additional strong assumptions. Demler *et al.*⁸ predicted that the correlation length for the superconducting proximity effect across an AF phase should be very long, and hence supercurrent should flow through a thick barrier in HTS/AF/HTS trilayer junctions. Their derivation was based on the SO(5) model⁷, but their conclusion is in fact more general: as long as HTS and AF phases are nearly degenerate, the proximity effect ought to be strong.

For the present study, we synthesized a number of superconductor–insulator–superconductor (SIS) heterostructures by stacking integer numbers of layers of $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ (LSCO) and La_2CuO_4 (LCO) in various sequences, using the atomic-layer synthesis capability of the advanced molecular beam epitaxy (MBE) system at Oxxel⁹. The films were essentially atomically smooth and perfect. For *c*-axis transport measurements, the films were patterned into

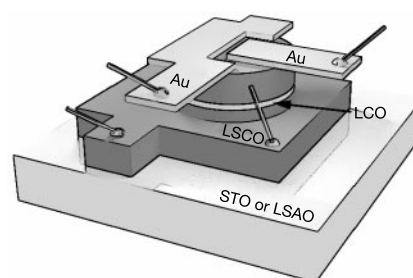


Figure 1 The structure of the SIS ‘sandwich’ junctions studied here. Either LaSrAlO_4 (LSAO) or SrTiO_3 (STO) is used for the substrate. The top and the bottom HTS electrodes are made of LSCO. The barrier, only 1UC thick, is made of LCO, which can be switched between insulating and HTS states by annealing in vacuum or in ozone, respectively, without affecting the HTS LSCO electrodes¹⁰. Two Au contacts to the top LSCO electrode and two to the bottom one allow for four-point contact transport measurements. Film perfection was ensured by reflection high-energy electron diffraction (RHEED) monitoring *in-situ*, and subsequent detailed characterization by X-ray diffraction (XRD) and AFM¹⁰. XRD showed very high crystallinity, absence of any secondary phases, rocking curves with full-width at half-maximum of 0.1° (indicating that the crystalline coherence length was equal to the film thickness), and pronounced low-angle reflectance oscillations and finite thickness oscillations. AFM showed terraces and one-unit-cell-tall steps that originate in a slight miscut of the substrate from the ideal (001) orientation; the r.m.s. surface roughness over an area of 2,500 μm^2 was 0.2–0.5 nm, much less than the height of a single unit cell. A perfect top surface indicates equally smooth multilayer interfaces, all of which were formed *in situ*; this was critical for the success of the experiment.

devices with cylindrical mesa structures (Fig. 1). On each chip, we fabricated a number of such devices to test the uniformity and, by varying the mesa diameter from 10 to 80 μm , the scaling of device parameters (such as the critical current I_c and the normal-state resistance R_n) with the mesa cross-section.

In Fig. 2, we show resistance–temperature curves for several sets of devices on the same chip. In each device, the barrier consists of a one-unit-cell (1UC)-thick layer of LCO, so its thickness is fixed at $d = 1.3$ nm. Our central result is that we observed no supercurrent, even with the lowest current bias (1 μA) and at the lowest temperature (4.2 K) accessible to us—unless LCO was made superconductive itself by annealing in ozone, or unless there was an obvious growth defect (a secondary-phase grain, or an a -axis-oriented grain) visible in atomic force microscopy (AFM) or even under an optical microscope. (Such a defect occurred in 3 devices (not shown) out of 36 on this particular chip—this allowed us to measure directly the resistance above the superconducting transition temperature (T_c) of the LSCO electrodes and leads.) We have synthesized, processed and tested dozens of other films and hundreds of devices, varying d from 1.3 to 20 nm (that is, from 1UC to 15UC of LCO) and saw the same result. We made a dozen LSCO:LCO superlattices as well. The findings were fully consistent with the results shown here, in most cases quantitatively.

In addition, we annealed a number of films and devices in vacuum or in O_2 (at low pressure, or at high pressure up to 200 bar). This modified LCO somewhat, but it stayed insulating¹⁰, and our central finding—the complete blockage of supercurrent by a 1UC layer of LCO—was not affected. In contrast, annealing in ozone

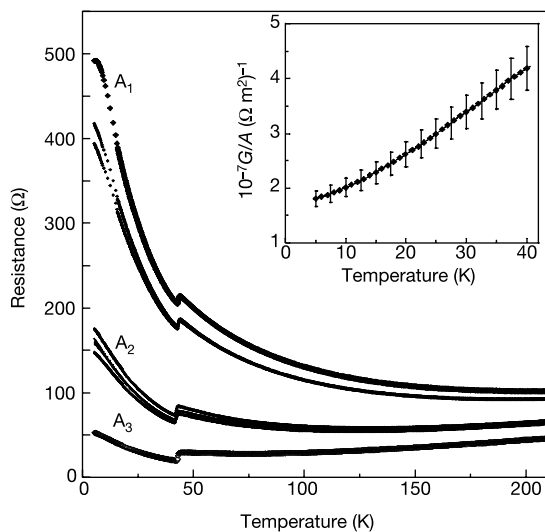


Figure 2 Transport in SIS junctions. The LCO barriers are 1UC thick in all the junctions shown here. The mesa cross-section varies from one set of curves to another: $A_1 = 1.8 \times 10^{-6} \text{ cm}^2$, $A_2 = 3.1 \times 10^{-6} \text{ cm}^2$ and $A_3 = 7.1 \times 10^{-6} \text{ cm}^2$. The figure illustrates good uniformity (the resistance is nearly the same, within $\sim 10\%$, in devices of nominally the same diameter) and approximate scaling of the resistance with the inverse of the area. A dip at 42 K shows where the LSCO electrodes undergo the superconducting transition; below this temperature, the voltage drops only across the LCO barrier. We do not have a direct check on magnetic ordering in our ultrathin LCO layers; however, bulk LCO samples with comparable resistivity (and similar temperature dependence) are indeed AF, with the ordering (Néel) temperature $T_N > 200$ K. Inset: average conductance per unit area, for $T < T_c$. A reasonable fit is obtained by assuming that $G(T) = G_0 + G_N(T)$, and $G_N(T) = A \log(T_0/T)$, a functional form registered earlier in bulk LSCO in a strong magnetic field²². This fitting procedure is not unambiguous, because of the limited temperature range (4.2–300 K) available to us; other fits are certainly possible; for example, to the Anderson–Zou law²³ $\rho = A/T + BT$, or to a power law such as that expected for phonon-assisted hopping through localized states^{11,12}. But in any case, the extrapolated G_0 is very small, as expected from a high barrier for tunnelling.

makes LCO superconductive, which is useful in control experiments. In this way, we get SS'S structures (where both S and S' are superconductive, but the critical temperature $T'_c < T_c$), and at $T = 4.2$ K these sustain the supercurrent density $j_c \approx 5 \times 10^4 \text{ A cm}^{-2}$, essentially the same as $j_c^\perp$ in single-phase LSCO films (that is, without any barrier). After re-annealing in vacuum or in O_2 , the LCO barrier reverts to insulating behaviour, and reduces the supercurrent to less than our detection limit, 10^{-6} A, even in the largest of our devices, 80 μm in diameter. This corresponds to $j_c < 2 \times 10^{-2} \text{ A cm}^{-2}$ —a reduction by a factor of at least 2.5×10^6 .

To estimate the tunnelling decay length x_0 , we use a simple one-dimensional (1D) potential barrier model: $T = C \exp(-d/x_0)$, where T is the barrier transmittance, $C = 16k_F^2\alpha^2/(k_F^2 + \alpha^2)^2$ is a number of the order of one (to be more precise, $C \approx 3$ for the Fermi energy $E_F \approx 0.3$ eV), $k_F^2 = 2mE_F/\hbar^2$, $\alpha^2 = 2m(V_0 - E_F)/\hbar^2$, m is the electron mass, V_0 is the barrier height, $x_0 = 1/2\alpha$, and d is the barrier thickness. This approximation is reasonable because of the atomic flatness of the films and because of the huge diameter-to-thickness ratio (10,000:1 or more), which makes the tunnelling direction very well defined in our devices—in contrast to tunnelling from the tips of scanning tunnelling microscopes (STMs), or through point-contact, step-edge, or bi-crystal junctions. Here, the supercurrent (if any) should flow in the c -axis direction.

By virtue of our digital deposition method, we know the barrier thickness, $d = 1.3$ nm. But because the barrier transmission T is exponentially sensitive in d , we checked for the possibility that one or more LSCO layers next to the LCO barrier could be carrier-depleted and non-superconductive, which would make the barrier effectively thicker. To this end, we synthesized several $(1 \times \text{LSCO}): (n \times \text{LCO})$ superlattices, where 1UC-thick layers of LSCO were interstacked with thicker layers of LCO, and measured their in-plane transport. If the superlattice film is superconductive, it means that even the very first LSCO layer next to LCO remains superconductive. In Fig. 3, we show the susceptibility of such a superlattice film after low-temperature annealing in oxygen—a treatment that leaves LCO insulating, as stated above. The film was apparently superconductive at 4.2 K (the temperature at which we compared I_c with and without the barrier) and even far above. We conclude that the barrier was indeed limited to a 1UC layer of

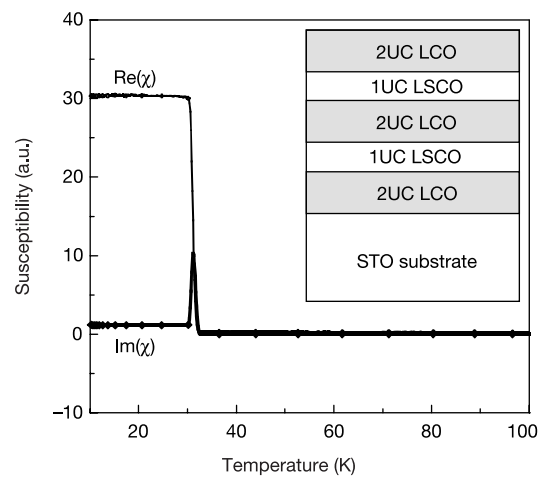


Figure 3 A 'reverse' experiment showing that 1UC-thick layers of LSCO remain superconducting. A superlattice film was synthesized by alternating 2UC-thick layers of insulating LCO and 1UC-thick layers of LSCO, as illustrated schematically in the inset; the actual film has 20 super-periods, and its total thickness is 79 nm. The temperature dependence of its susceptibility (the main panel), measured by the mutual inductance technique, shows a sharp superconducting transition. T_c is somewhat reduced here because we used a different substrate, STO, while the devices shown in Fig. 2 were grown on LSAO; for a discussion of substrate effects, see ref. 10 and references therein.

LCO, that is, that $d = 1.3$ nm. The barrier height we do not know *a priori*; for $d = 1.3$ nm, our measured value of $T \approx 10^{-6}$ corresponds to $x_0 \approx 1$ Å and $V_0 - E_F \approx 1$ eV.

We can reinforce this conclusion by showing that it also follows from our normal-state transport data. In the inset to Fig. 2, we plot the average conductance $G(T)$ normalized to unit area, for $T < T_c$. It is apparent that the devices are 'intrinsically shunted' by some kind of temperature-dependent hopping process; this is indeed typical of HTS tunnel junctions^{11,12}. We write $G(T) = G_0 + G_h(T)$ and separate the elastic (temperature independent) contribution G_0 from the hopping term $G_h(T)$ by fitting and extrapolating to zero temperature. However, it is clear that the extra contribution G_0 from direct tunnelling must be small, as the total $G(T)$ is similar to what we see in single-phase LCO films prepared under similar conditions, and also to what is seen in bulk single crystals. From Fig. 2 inset, we see that the extrapolated $G_0/A \approx (1-2) \times 10^7$ ($\Omega \text{ m}^2$)⁻¹, which for our junction size indeed corresponds to a very small G_0 . Using the same 1D model as above¹³, from $G_0/A = (1/4\pi)(e^2/h)(1/dx_0)\exp(-d/x_0)$ we again get $x_0 \approx 1$ Å, which corresponds to $V_0 - E_F \approx 1$ eV and $T \approx 10^{-6}$, consistent with our inference from supercurrent data. (A more rigorous approach (A. Leggett, personal communication) is to use the Landauer formula $G = (2e^2/h)TN$, where N is the number of transverse modes, and e is the electron charge; this gives $1/T = (4-8) \times 10^6$, not too different from our simple model.)

We now draw some conclusions from our work. First, for Josephson junction electronics these devices are of little practical value; the barrier is too strong. Second, the absence of pinholes and superconducting shorts even in IUC barriers attests to excellent control of the growth process. Third, the MHI and HTS states seem to be remote, separated by about 1 eV per electron. This is a very robust conclusion that essentially follows from the raw data by looking at the orders of magnitude. The two phases do not mix; the HTS/MHI interface is rather sharp, on the length scale of 1 Å. In our experiments, this manifests itself in two ways: as the absence of supercurrent through ultrathin LCO barriers, which indicates that the superconducting order parameter does not 'leak' into the insulator, and as the persistence of HTS in ultrathin LSCO layers (within the LSCO:LCO superlattice structures), which shows that on the HTS side the order parameter is preserved right up to the interface. (HTS also persists in half-unit-cell-thick layers of Bi₂Sr₂CaCu₂O₈; ref. 14.) It remains to be seen whether (a variant of) the Hubbard model can be reconciled with the present experimental findings.

We now consider the influence of anisotropy on our results. In LSCO, the coupling along the c axis is much weaker than along the a or b axes—using j_c as the yardstick, by as much as a factor of 500 or so. Note, however, that we have taken this factor into account already, by comparing j_c in our devices to the c -axis j_c of LSCO, which our IUC barrier cuts down by an additional factor of 10^6 . Hence, the absence of proximity effects cannot be simply attributed to the geometry of the experiment.

Our data also impose some constraints on the nature of states created by doping. In the conventional 'doped semiconductor' (or 'rigid band') picture, one assumes that the valence band is fully occupied in the insulating LCO, and that it gets partly vacated from the top when holes are introduced by doping. In this case, one would expect μ to undergo a large shift (~ 1 V) on doping, with consequent formation of a substantial depletion layer. However, the absence of a major reduction in T_c even in the first LSCO layer next to the insulating LCO layer implies that the hole density did not change much: $\Delta T_c \approx 5-10$ K corresponds to $\Delta T_c/T_c \approx 10-20\%$ and (as T_c is approximately proportional to the hole density n and the Sr doping level x) to $\Delta n/n \approx \Delta x/x \approx 10-20\%$, and since $x \approx 0.15$, $\Delta x < 0.03$, and hence $e\Delta\mu < 0.1$ eV (M. Beasley, personal communication). The same conclusion indeed follows from the fact that IUC LCO layers stay insulating next to the metallic LSCO.

Another possibility is that doping creates new states inside the gap. In undoped copper oxides such as LCO the optical gap is about 2 eV wide, according to a variety of experiments (such as optical absorption, photoconductivity, resonant Raman scattering, luminescence)¹⁵⁻¹⁷. If the doped states are at the midgap, the LSCO/LCO barrier should be about 1 eV high, and one would expect the change in μ on doping to be small, with virtual absence of a depletion layer at the HTS/MHI interface. Our data clearly favour this picture.

Our experiments with ultrathin LCO and LSCO layers show an absence of proximity effects between HTS and MHI phases, and indicate that doped electronic states appear at the midgap. This is consistent with the model that postulates spin-charge separation^{2,3}. Alternative explanations may be offered^{4-6,18-21}, but more theoretical work may be needed to determine which of these are compatible with the constraints imposed by the present experimental findings. To discriminate further between the alternatives, a direct experimental measurement of the spin of free charge carriers would be invaluable. □

Received 8 October 2002; accepted 25 February 2003; doi:10.1038/nature01544.

- Anderson, P. W. *The Theory of Superconductivity in the High- T_c Cuprate Superconductors* (Princeton Univ. Press, Princeton, 1997).
- Anderson, P. W. Spin-charge separation is the key to the high T_c cuprates. *Physica C* **341-348**, 9-10 (2000).
- Kivelson, S. A., Rokhsar, D. S. & Sethna, J. P. Topology of the resonating valence-bond state: Solitons and high- T_c superconductivity. *Phys. Rev. B* **35**, 8865-8868 (1987).
- Zaanen, J. & Gunnarsson, O. Charged magnetic domain lines and the magnetism of high- T_c oxides. *Phys. Rev. B* **40**, 7391-7394 (1989).
- Fisher, D. S., Kotliar, G. & Moeller, G. Midgap states in doped Mott insulators in infinite dimensions. *Phys. Rev. B* **52**, 17112-17118 (1995).
- Kampf, A. & Schrieffer, J. R. Pseudogaps and the spin-bag approach to high- T_c superconductivity. *Phys. Rev. B* **41**, 6399-6408 (1990).
- Zhang, S. C. A unified theory based on SO(5) symmetry of superconductivity and antiferromagnetism. *Science* **275**, 1089-1096 (1997).
- Demler, E., Berlinsky, A. J., Kallin, C., Arnold, G. B. & Beasley, M. R. Proximity effect and Josephson coupling in the SO(5) theory of high- T_c superconductivity. *Phys. Rev. Lett.* **80**, 2917-2920 (1998).
- Bozovic, I. Atomic-layer engineering of superconducting oxides: Yesterday, today, tomorrow. *IEEE Trans. Appl. Supercond.* **11**, 2686-2695 (2001).
- Bozovic, I., Logvenov, G., Belca, I., Narimbetov, B. & Svelko, I. Epitaxial strain and superconductivity in La_{2-x}Sr_xCuO₄ thin films. *Phys. Rev. Lett.* **89**, 100700 (2002).
- Golubov, A. A. *et al.* Resonant tunneling in Y(Dy)Ba₂Cu₃O_{7-δ}/PrBa₂Cu_{3-x}Ga_xO_{7-δ}/Y(Dy)Ba₂Cu₃O_{7-δ} ramp-type Josephson junctions. *Physica C* **235-240**, 3261-3262 (1994).
- Bozovic, I. & Eckstein, J. N. Transport in atomically engineered BiSrCaCuO multilayers. *J. Supercond.* **8**, 537-540 (1995).
- Wolf, E. L. *Principles of Electron Tunneling Spectroscopy* 36-45 (Oxford Univ. Press, Oxford, 1989).
- Bozovic, I., Eckstein, J. N., Klausmeier-Brown, M. E. & Virshup, G. F. Superconductivity in epitaxial Bi₂Sr₂CuO₆/Bi₂Sr₂CaCu₂O₈ superlattices: The superconducting behavior of ultrathin cuprate slabs. *J. Supercond.* **5**, 19-23 (1992).
- Thio, T., Birgeneau, R. J., Cassanho, A. & Kastner, M. A. Determination of the energy gap for charged excitations in insulating La₂CuO₄. *Phys. Rev. B* **42**, 10800-10803 (1990).
- Uchida, S. *et al.* Optical spectra of La_{2-x}Sr_xCuO₄: Effect of carrier doping on the electronic structure of the CuO₂ plane. *Phys. Rev. B* **43**, 7942-7954 (1991).
- Kirillov, D., Bozovic, I., Char, K. & Kapitulin, A. Resonance effects in Raman scattering in YBa₂Cu₃O_{7-x}. *J. Appl. Phys.* **66**, 977-979 (1989).
- Gor'kov, L. P. & Sokol, A. V. Phase stratification of an electron liquid in the new superconductors. *JETP Lett.* **46**, 420-423 (1987).
- Emery, V. J. & Kivelson, S. A. Frustrated electronic phase separation and high-temperature superconductors. *Physica C* **209**, 597-621 (1993).
- Alexandrov, A. S. & Mott, N. *Polarons and Bipolarons* (World Scientific, Singapore, 1995).
- Varma, C. M., Littlewood, P. B., Schmitt-Rink, S., Abrahams, E. & Ruckenstein, A. E. Phenomenology of the normal state of Cu-O high-temperature superconductors. *Phys. Rev. Lett.* **63**, 1996-1999 (1989); erratum: *Phys. Rev. Lett.* **64**, 497 (1990).
- Ando, Y., Boeinger, G. S., Passner, A., Kimura, T. & Kishio, K. Logarithmic divergence of both in-plane and out-of-plane normal-state resistivities of superconducting La_{2-x}Sr_xCuO₄ in the zero-temperature limit. *Phys. Rev. Lett.* **75**, 4662-4665 (1995).
- Anderson, P. W. & Zou, Z. "Normal" tunneling and "normal" transport: Diagnostics for the resonating-valence-bond state. *Phys. Rev. Lett.* **60**, 132-135 (1988).

Acknowledgements We acknowledge discussions with M. R. Beasley, A. Leggett, G. A. Sawatzky, S. Kivelson, A. Tsvetkov, P. Stamp, S. C. Zhang, A. Kleinsasser, A. Abrikosov, A. Millis, R. Dynes and S. Chakraverty. We thank I. Svelko, B. Narimbetov and I. Belca for technical help. This work was supported in part by AFOSR.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to I.B. (e-mail: ibozovic@pacbell.net).

Experimental detection of α -particles from the radioactive decay of natural bismuth

Pierre de Marcillac, Noël Coron, Gérard Dambier, Jacques Leblanc & Jean-Pierre Moalic

Institut d'Astrophysique Spatiale, CNRS & Université Paris Sud, UMR 8617, Bât. 121, 91405 Orsay Cedex, France

The only naturally occurring isotope of bismuth, ^{209}Bi , is commonly regarded as the heaviest stable isotope. But like most other heavy nuclei abundant in nature and characterized by an exceptionally long lifetime, it is metastable with respect to α -decay¹. However, the decay usually evades observation because the nuclear structure^{2,3} of ^{209}Bi gives rise to an extremely low decay probability and, moreover, generates low-energy α -particles difficult to detect. Indeed, dedicated experiments^{2–6} attempting to record the α -decay of ^{209}Bi in nuclear emulsions failed. However, scintillating bolometers^{7–9} operated at temperatures below 100 mK offer improved detection efficiency and sensitivity, whereas a broad palette of targets could be available¹⁰. Here we report the successful use of this method for the unambiguous detection of ^{209}Bi α -decay in bismuth germanate detectors cooled to 20 mK. We measure an energy release of $3,137 \pm 1$ (statistical) ± 2 (systematic) keV and a half-life of $(1.9 \pm 0.2) \times 10^{19}$ yr, which are in agreement with expected values.

Mass excess measurements show that ^{209}Bi , the only naturally abundant isotope of bismuth, should decay to the more stable thallium isotope ^{205}Tl (Fig. 1a,b). The low available decay energy Q_α induces a vanishing probability P_α for tunnelling of the α -particle through the nuclear potential barrier. Moreover, ^{209}Bi is next to doubly 'magic' nucleus ^{208}Pb ($Z = 82$; $N = 126$) and the need to pick nucleons across this extremely stable configuration, together with the poor overlap of initial and final nuclear states wave functions, further explain why the α -decay of ^{209}Bi is very rare, and thus difficult to detect.

Direct observation of the decay is also not easy: the α -particles move in solid matter over distances of around $10 \mu\text{m}$ (ref. 11), which are impractical to detect using common detectors (gas counters, semiconductors). It was thought that the appearance in 1945 of commercial nuclear emulsions that could be loaded with Bi would solve the technical issues of detection. Numerous experiments followed^{2–6}: the exposure times were long, and those results found positive^{5,6} were invalidated afterwards (Table 1).

We detected α -decay of ^{209}Bi using the scintillating bolometer technique. The earlier scintillating bolometers used $\text{CaF}_2(\text{Eu})$ and a silicon photodiode, glued on the crystal, as the light detector². After recording an initial high-resolution α spectrum within a micro-target (a 1-mg diamond bolometer at 1.3 K, ref. 12), we introduced a 'heat-and-light' pair of bolometers, with 300 mg of scintillating $\text{CaF}_2(\text{Eu})$ at 130 mK (ref. 8). As part of our development of more massive scintillating bolometers (Fig. 1c) for the Dark Matter search experiment ROSEBUD, a 45.7-g BGO ($\text{Bi}_4\text{Ge}_3\text{O}_{12}$) bolometer exhibiting unusual discriminating properties at 20 mK was checked at Orsay for contamination at high energy (Fig. 2): this crystal had already shown a strong ^{207}Bi contamination¹³. Thanks to their high quenching factor, α -particles could be perfectly separated from events induced by cosmic rays (see Supplementary Information). An unexpected α line was observed at $E = 3,130 \pm 16$ keV, now attributed to ^{209}Bi decay. The publicly available SRIM-2003 package can be used to estimate the ranges in BGO of the decay products—respectively $7.9 \mu\text{m}$ for the α -particle ($E_\alpha = 3,077$ keV) and about $170 \text{ \AA}$ for the ^{205}Tl recoiling nucleus ($E_{\text{Tl}} = 60$ keV). These events

are fully contained in our massive crystal, guaranteeing 100% detection efficiency. The associated half-life of ^{209}Bi is then calculated from the events rate to be $T_{1/2} = 1.44\text{--}1.95 \times 10^{19}$ years (68% confidence level, CL).

In all materials tested, bolometers register indiscriminately α -particles and their associated nuclear recoils with nearly equal

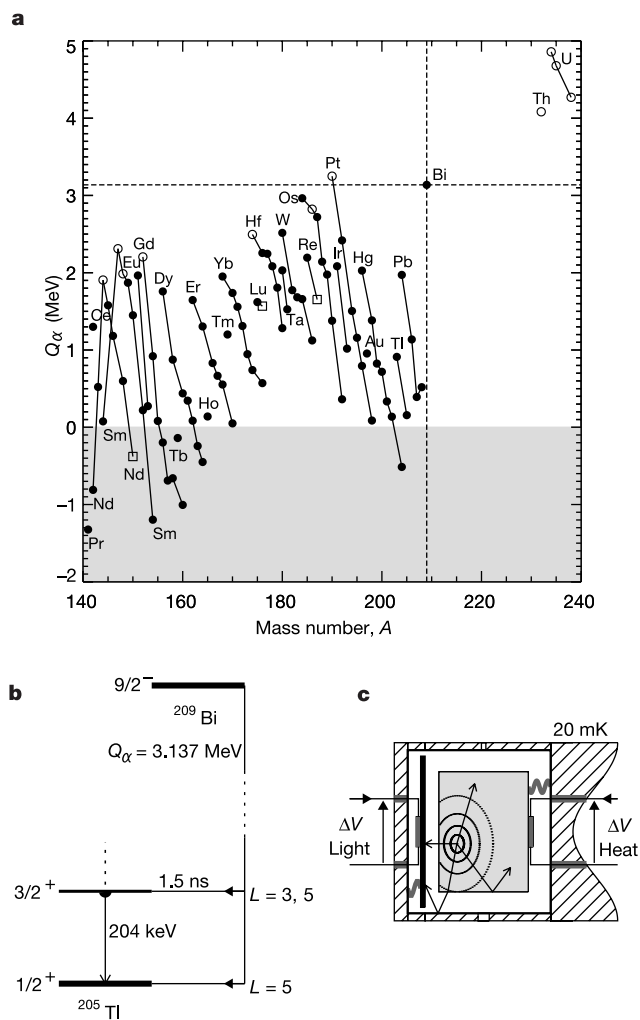


Figure 1 The bolometric detection of ^{209}Bi decay. **a**, Stability of heavy ($A > 140$) naturally abundant elements. The mass excess Q_α , calculated from tables¹, is given for elements found in average abundance $>0.1 \text{ mg ton}^{-1}$ in the Earth's crust. Lines connect isotopes of the same element. α -particle decay is energetically allowed only for $Q_\alpha > 0$ (out of the grey area). Squares, β -decay reported; black circles α -decay not yet observed; white circles, α -decay detected. Element (year of report): ^{144}Nd (1954); ^{147}Sm (1933); ^{148}Sm (1970); ^{152}Gd (1959); ^{174}Hf (1959); ^{186}Os (1975); ^{190}Pt (1921); ^{232}Th (1898); $^{234,235,238}\text{U}$ (1896). **b**, α -decay scheme of ^{209}Bi . Nuclear states are shown with their adopted spin and parity (J^π), energy level and half-life²⁷. The finite quanta of angular momentum L , to be carried by the α -particle according to the transition rules, are indicated. **c**, Measuring heat and light in scintillating bolometers. Scintillating bolometer (BGO cylinder; diameter = height = 20.2 mm) and optical bolometer (germanium disk; diameter = 25 mm; 100 μm thick) face each other in an Ag-coated light-reflecting cavity. A continuous base temperature $T = 20 \text{ mK}$ —the key issue for high sensitivity—is provided by a ^3He – ^4He dilution refrigerator, to which the detectors are thermally linked. Thermo-resistive sensors made from neutron transmuted doped germanium (NTD-Ge) are polarized at constant current and convert temperature rise into voltage signal. The voltage pulses (μV) that follow an interaction are amplified, digitalized and stored (see examples in Supplementary Information). Suspensions are not drawn. Small holes in the cavity allow external excitation, either with optical pulses sent through optical fibres, or with external radioactive sources.

Table 1 Survey of attempts to detect ^{209}Bi decay

Year of publication	Final level tested in ^{205}Tl	Detection technique	Effective exposure*		Number of reported (expected) events†
			Atoms $\times$ year	Time	
1949 (ref. 4)	Ground state	N.E.	4.3×10^{15}	≤ 7 months	< 1 (0)
1951 (ref. 5)	Ground state	N.E.	7.0×10^{18}	2.25 years	18 (0)
1952 (ref. 6)	Ground state	N.E.	2.0×10^{18}	3.3 months	7 (0)
1958 (ref. 2)	Ground state	N.E.	2.9×10^{18}	3 years	< 1 (0)
1972 (ref. 3)	Ground state	N.E.	2.3×10^{20}	"years..."	< 1 (9)
2000 (ref. 20)	204 keV	$\text{Bi}_2\text{O}_3 + \gamma$ spectrometer	3.0×10^{21}	29.3 days	< 46 (1–3)
2003 (this work)	Ground state	BGO bolometers	3.5×10^{21}	5 days	128
2003 (this work)	204 keV	46-g BGO bolometer	1.1×10^{20}	5 days	≤ 1 (0)

N.E.: Bi-loaded Nuclear Emulsion (Ilford type C2^{2,4,6} or L4³); the energy of the α -particles is inferred from their range in photographic plates, with typical energy resolution 700 keV FWHM²⁶.

*Effective exposure (atoms $\times$ years) includes authors' own estimation of the detection yield, taking into account microscopist's efficiency, background level, absorption of gammas and so on. To obtain the published limit, or value, of $T_{1/2}$, multiply this number by $0.693 = \ln(2)$ and divide by the number of reported events.

†The expected number of events is deduced from this work. The false-positive events reported in nuclear emulsions^{5,6} might result from "chemical fog"²⁶, while, on the other hand, uncontrolled fading of plates could explain why detection was marginally overlooked in the longest exposure time experiment³. Updated data from the Oroville's experiment²⁰ were kindly provided by those authors.

thermal conversion yields^{14,15} (see also discussion in Supplementary Information). Assuming the equality condition to be true also for BGO, we mounted and tested a second BGO bolometer (91.2 g), under external ^{241}Am irradiation (Fig. 3). The crystal was grown from purer raw material (Crismatec). Although its ^{207}Bi content was in fact strongly reduced (see spectrum in Supplementary Information), the former α line was still observed at a rate compatible with mass scaling, $T_{1/2} = 1.81\text{--}2.29 \times 10^{19}$ years (68% CL).

Combining the previous half-lives measurements increases statistical significance, to give $T_{1/2}(^{209}\text{Bi}) = 1.9 \pm 0.2 \times 10^{19}$ years (68% CL), which we compared with the few published theoretical expectations^{2,3}. We followed the later estimation³ in which, as suggested by Rasmussen, the ^{209}Bi decay rate was calculated differentially from the decay rate of ^{210}Po ($T_{1/2} = 138.4$ days), on the basis of their nuclear structure similarities. After updating the relevant atomic masses and energy values, we revised the tabulated barrier penetration probabilities P_α (ref. 16,17), including the

'centrifugal potential' attributable to the finite angular momentum, and found an expected half-life $T_{1/2} = 4.6 \times 10^{19}$ years (formerly 1.1×10^{20} years) in close agreement with our experimental result.

The curvature of the α branch in Fig. 3, which showed increasing light emission with energy, implied at least some intrinsic non-linearity on the thermal channel as well, owing to conservation of energy. A slight opposite trend, which was indeed found on the heat channel, was easily corrected, as an accurate calibration through the well-identified α lines reduced the integral non-linearity below 10^{-3} over the whole spectrum (0–9 MeV). The energy of the new line, shifted by -12 keV in this process, was then calculated as $E = 3,137 \pm 1$ (stat.) ± 2 (syst.) keV, precisely matching the

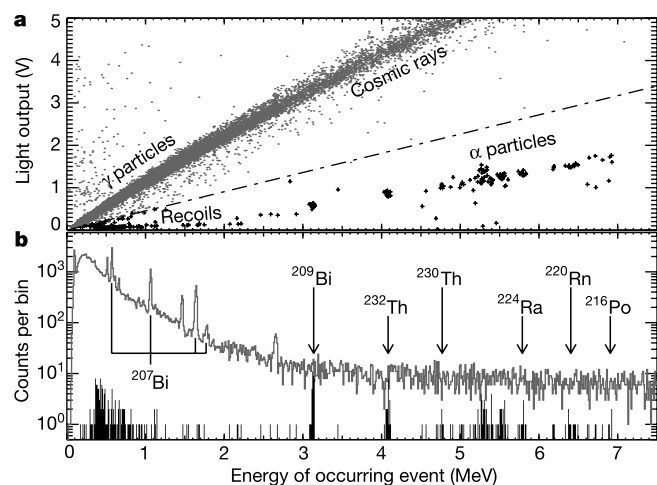


Figure 2 Discrimination of background events in a 46-g BGO bolometer. **a**, Background in the 46-g BGO bolometer (100 h). 'Light' signal versus 'Heat' signal. The latter has been converted into energy of occurring event in the BGO crystal. Two main classes of events are visible: a ' γ and cosmic rays' branch, giving high light signals owing to the high ionization power of these particles, and an α -particle branch. A third class of events is visible at the low-energy, low-light level: it is linked to nuclear recoils induced by ambient neutrons at laboratory level¹³. Separate calibration is needed^{13,22}: ^{207}Bi lines from internal contamination are used for the γ branch; for the α branch, internal ^{232}Th , ^{230}Th , ^{224}Ra , ^{220}Rn and ^{216}Po α lines were identified step by step. Some were obviously paired events (see Supplementary Information); the final α calibration uses a weighted least-squares polynomial fit to these lines, including the 'zero energy' point. **b**, Energy spectra of alphas, recoils and gammas after application of a simple cut on light-to-heat signal ratio (dotted line in **a**). The expected location of all well-identified lines is reported, the γ spectrum being dominated by lines from ^{207}Bi ($T_{1/2} = 35.5$ yr; 3 Bq kg^{-1})¹³.

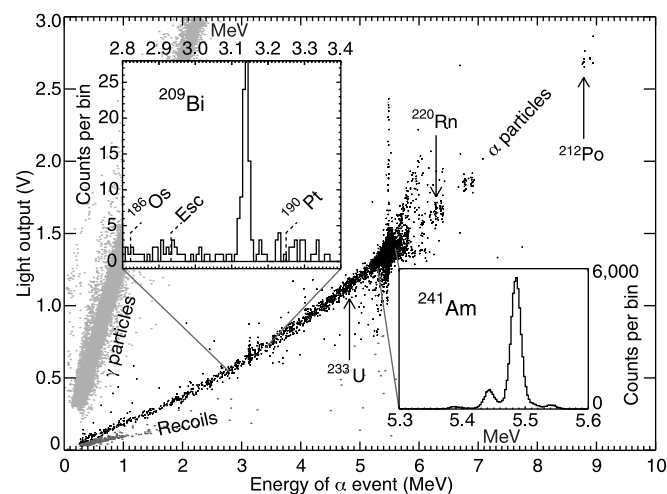


Figure 3 α events during ^{241}Am calibration in a 91-g BGO bolometer (114 h). α events (black dots) are selected according to their light-to-heat ratio (see Fig. 2). External ^{241}Am is used for calibration: the source is thin (activity $\approx 185 \text{ Bq cm}^{-2}$), and negligible energy loss in the source itself is assumed (see Supplementary Information). A low-rate, low-energy tail is visible, associated to the (back)scattering of α -particles in the support: these events allow for indisputable assessment of the newly detected line to α -decay, a unique feature of this technique. Events apart from the ^{241}Am lines were shown to be slight contaminations of the source itself: the associated events from ^{233}U , ^{220}Rn and ^{212}Po were used, together with the ^{241}Am and 'zero energy' lines, to provide, after a weighted least-squares polynomial fit, an accurate calibration of occurring α events (see details in Supplementary Information); their expected location is reported on the diagram. Right inset, spectrum of ^{241}Am . The fine structure is resolved, owing to the 18 keV FWHM resolution of the detector (the noise measured on the baseline has a negligible 6-keV contribution to this width). The same spectrum looked at from the light channel would show a 295-keV FWHM resolution only. Left inset, the ^{209}Bi α -decay line. Spectrum of recorded events (2.8 MeV to 3.4 MeV) during the 114-h ^{241}Am calibration run. The ^{209}Bi line stands out well from the background (around seven events estimated under the line). The location of conceivable α -decays in this energy window is indicated: ^{186}Os , ^{190}Pt , and ^{209}Bi decay to ^{205}Tl 204 keV level γ escape peak (Esc).

expected value of ^{209}Bi α -decay, $Q_\alpha = 3,137 \pm 0.9 \text{ keV}$ (ref. 1, see comment in Supplementary Information).

Finally, possible contaminations by already-known α emitters of close Q_α were considered: ^{186}Os ($T_{1/2} = 2.0 \times 10^{15}$ years; $Q_\alpha = 2,822 \text{ keV}$) and ^{190}Pt ($T_{1/2} = 6.5 \times 10^{11}$ years; $Q_\alpha = 3,249 \text{ keV}$) were discarded because the results implied very high atomic concentrations, respectively Os/Bi $\approx 6,500 \text{ p.p.m.}$ and Pt/Bi $\approx 250 \text{ p.p.m.}$ A dedicated quantitative Pt implantation experiment in BGO, followed by a secondary ion mass spectrometry (SIMS) analysis, proved that Pt concentration was under 2 p.p.m. Higher concentrations were very unlikely¹⁸, and would have strongly affected the scintillating properties of the crystal¹⁹.

The auxiliary observation of ^{209}Bi decay to the first excited state of ^{205}Tl (Fig. 1b) would constitute a conclusive test of our proper identification of the α line. Such a—negative—search was performed recently, using a very low background facility under the Oroville dam, where 204 keV γ -particles emerging from 1.8 kg of Bi_2O_3 were looked for²⁰. Thanks to the high Z of Bi, and to their slow time constant ($\tau \approx \text{a few ms} \gg 1.5 \text{ ns}$), BGO bolometers naturally sum the energies of the α -particle and of the following absorbed γ , adding another event to the ‘full energy’ peak. However, the γ escape probability is not negligible: we estimate it to be 12.8% in our 46-g BGO detector by Monte Carlo simulations. Taking into account internal conversion²¹, the overall efficiency for detecting α -decay at this level through recording of the escape peak is $\eta = 8.8\%$. With reasonable assumptions (see Supplementary Information), the associated partial half-life is estimated to be $T_{1/2} = 6\text{--}13 \times 10^{20}$ years: a two-month exposure would then probably lead to its detection (Table 1).

These data extend by more than three orders in magnitude the range of α -particle decays ever reported ($T_{1/2} = 7 \times 10^{15}$ years; ^{148}Sm). α -particle events may now be accurately tracked down to 500 keV , with background suppression efficiency over 98% , in common laboratory radioactive environments. Equivalent rejection power against γ -particles should be observed down to around 50 keV , provided that the confusing recoils background is reduced (in underground experiments, for example; see Supplementary Information). These possibilities in the search for rare events—indicated also by our recent improvements of the ^{180}W α -particle decay rate limit²²—justify the greater complexity added by the low-temperature requirements.

With this ^{209}Bi decay detection, BGO bolometers open up the rarest of natural main energy lines—whereas much rarer double β -particle decays are observed among naturally abundant elements²³, no line corresponding to the neutrinoless mode $0\nu\beta\beta$ has been hitherto convincingly reported. This particular property motivated recent searches for ^{209}Bi decay: with respect to the search for proton decay, a proposal was made to use ‘freshly produced’ bismuth to check for deviation of radioactive decay from the exponential law at small times^{24,25}. □

Received 20 November 2002; accepted 10 March 2003; doi:10.1038/nature01541.

- Audi, G., Bersillon, O., Blachot, J. & Wapstra, A. H. The Nubase evaluation of nuclear and decay properties. *Nucl. Phys. A* **624**, 1–124 (1997).
- Hincks, E. P., Millar, C. H. & Hanna, G. C. A search for α -particles from the decay of ^{209}Bi . *Can. J. Phys.* **36**, 231–251 (1958).
- de Carvalho, H. G. & de Araújo Penna, M. Alpha-activity of ^{209}Bi . *Lett. Nuovo Cim.* **3**, 720–722 (1972).
- Jenkner, K. & Broda, E. Some upper limits of possible alpha-activity. *Nature* **164**, 412–413 (1949).
- Faraggi, H. & Berthelot, A. Sur la radioactivité alpha du bismuth naturel. *C. R. Acad. Sci.* **232**, 2093–2095 (1951).
- Riezler, W. & Porschen, W. Natürliche Radioaktivität von Wismut. *Z. Naturforsch.* **7a**, 634–635 (1952).
- Gonzalez-Mestres, L. & Perret-Gallix, D. Detection of low energy solar neutrinos and galactic dark matter with crystal scintillators. *Nucl. Instrum. Meth. A* **279**, 382–387 (1989).
- Bobin, C. *et al.* Alpha/gamma discrimination with a $\text{CaF}_2(\text{Eu})$ target bolometer optically coupled to a composite infrared bolometer. *Nucl. Instrum. Meth. A* **386**, 453–457 (1997).
- Alessandrello, A. *et al.* Development of a thermal scintillating detector for double beta decay of ^{48}Ca . *Nucl. Phys. B (Proc. Suppl.)* **28A**, 233–235 (1992).
- Moses, W. W. *et al.* Prospects for dense, infrared emitting scintillators. *IEEE Trans. Nucl. Sci.* **45** (3), 462–466 (1998).
- Ziegler, J. F. *The Stopping and Ranges of Ions in Matter* Vol. 4, Helium: Stopping Powers and Ranges in All Elemental Matter (Pergamon, New York, 1977).

- Coron, N. *et al.* A composite bolometer as a charged-particle spectrometer. *Nature* **314**, 75–76 (1985).
- Cebrián, S. *et al.* The ROSEBUD experiment at Canfranc: 2001 report. *Nucl. Phys. B (Proc. Suppl.)* **110**, 97–99 (2002).
- Zhou, J. W. *et al.* Advances towards fast thermal detectors of intermediate mass with high resolution and large dynamic range. *Nucl. Instrum. Meth. A* **335**, 443–452 (1993).
- Alessandrello, A. *et al.* Preliminary results on the performance of a TeO_2 thermal detector in a search for direct interactions of WIMPs. *Phys. Lett. B* **384**, 316–322 (1996).
- Rasmussen, J. O. Alpha-decay barrier penetrabilities with an exponential nuclear potential: even-even nuclei. *Phys. Rev.* **113**, 1593–1598 (1959).
- Rasmussen, J. O. Alpha-decay barrier penetrabilities with an exponential nuclear potential: odd-mass nuclei. *Phys. Rev.* **115**, 1675–1679 (1959).
- Zhou, T. Q., Tan, H. R., He, C. F., Zhu, R. Y. & Newman, H. B. Determination of trace elements in BGO by neutron activation analysis. *Nucl. Instrum. Meth. A* **258**, 58–66 (1987).
- Barnes, R. G. L., Sims, R., Rousseau, M. D. & Sproston, M. Bismuth germanate (BGO) optimisation for energy resolution and purity. *IEEE Trans. Nucl. Sci.* **31**, 249–252 (1984).
- Norman, E. B. *et al.* Diamonds, maybe, but bismuth is not forever. *Bull. Am. Phys. Soc.* **45**, 30 (2000).
- Giannatiempo, A. & Perego, A. Penetration effects in the internal conversion process of the 204 KeV transition in ^{205}Tl . *Z. Phys. A* **308**, 247–251 (1982).
- Cebrián, S. *et al.* Improved limits for natural α radioactivity of tungsten with a CaWO_4 scintillating bolometer. *Phys. Lett. B* **556**, 14–20 (2003).
- Tretyak, V. I. & Zdesenko, Y. G. Tables of double beta decay data—an update. *Atom. Data Nucl. Data Tables* **80**, 83–116 (2002).
- DeBraeckeleer, L. *et al.* Measurement of lifetime of ^{209}Bi and test of the exponential decay law. *TUNL Progr. Rep.* **37**, 94 (1998).
- DeBraeckeleer, L., Gould, C. R. & Tornow, W. Search for $T_{1/2}$ of the “stable” nucleus ^{209}Bi . *TUNL Progr. Rep.* **38**, 87–89 (1999).
- Powell, C. F., Fowler, P. H. & Perkins, D. H. *The Study of Elementary Particles by the Photographic Method. An Account of the Principal Techniques and Discoveries Illustrated by an Atlas of Photomicrographs* 137–138, 644–645 (Pergamon, New York, 1959).
- Firestone, R. B. *Table of Isotopes*, 8th edn, Vol. II: A = 151–272 (eds Shirley, V. S., Baglin, C. M., Chu, S. Y. F. & Zipkin, J.) 2545, 2595 (Wiley, New York, 1996).

Supplementary Information accompanies the paper on Nature’s website (<http://www.nature.com/nature>).

Acknowledgements This work is supported by the R&D programme from CNRS/INSU, and by BNM for the high-resolution α -spectrometry part. The design of the optical detectors results from years of support by CNES. P. Pari from CEA/SPEC designed our dilution refrigerator. We thank the SEMIRAMIS members of CSNSM for their accurate implantation, as well as F. Jomard for SIMS analysis, G. Audi for discussions, E. Leblanc from LNH/BNM for independent calibration source measurements, and I. Rameau and Y. Bouvry for their support and comments.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for material should be addressed to P.d.M. (e-mail: pierre.demarcillac@ias.u-psud.fr) or N.C. (e-mail: noel.coron@ias.u-psud.fr).

Magmatic events can produce rapid changes in hydrothermal vent chemistry

Marvin D. Lilley*, David A. Butterfield†, John E. Lupton‡ & Eric J. Olson*

* School of Oceanography, University of Washington, Seattle, Washington 98195-5351, USA

† Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, and Pacific Marine Environmental Laboratory, NOAA, Seattle, Washington 98195, USA

‡ Pacific Marine Environmental Laboratory, NOAA, Newport, Oregon 97365, USA

The Endeavour segment of the Juan de Fuca ridge is host to one of the most vigorous hydrothermal areas found on the global mid-ocean-ridge system, with five separate vent fields located within 15 km along the top of the ridge segment¹. Over the past decade, the largest of these vent fields², the ‘Main Endeavour Field’, has exhibited a constant spatial gradient in temperature and chloride concentration in its vent fluids, apparently driven by differences in the nature and extent of subsurface phase separation³. This stable situation was disturbed on 8 June 1999 by an earthquake

swarm⁴. Owing to the nature of the seismic signals and the lack of new lava flows observed in the area during subsequent dives of the *Alvin* and *Jason* submersibles (August–September 1999), the event was interpreted to be tectonic in nature⁴. Here we show that chemical data from hydrothermal fluid samples collected in September 1999 and June 2000 strongly suggest that the event was instead volcanic in origin. Volatile data from this event and an earlier one at 9° N on the East Pacific Rise show that such magmatic events can have profound and rapid effects on fluid–mineral equilibria, phase separation, ³He/heat ratios and fluxes of volatiles from submarine hydrothermal systems.

There is no evidence of recent volcanism in the area of the Main Endeavour Field (MEF; Supplementary Fig. 1), and a 0.5-km-wide axial valley indicates that the area is currently in a tectonically controlled phase⁵. In 1988, a weak seismic reflector was imaged at about 2.5 km depth⁶, but a later refraction experiment showed no evidence for a low-seismic-velocity zone (magma chamber) beneath this reflector⁷. Recent seismic data indicate that on-axis micro-earthquakes persist to a depth of 3.5 km below the sea floor, well below the weak seismic reflector at 2.5 km, although most occur above 3 km (ref. 8). For these reasons, Endeavour has become the type example for a ‘cracking front’ hydrothermal system, in which heated sea water continually penetrates deeper within a frozen magma chamber and extracts heat from the gabbroic layer^{9,10}. However, new seismic reflection data collected in July 2002 indicate a magma body at about 2.3–2.6 km depth¹¹. This forces us to re-examine our ideas about the heat source driving the Endeavour hydrothermal system, and to consider whether periodic magma resupply¹² may be important.

Although the earthquake swarm was originally argued to be a tectonic event⁴, a reanalysis of the seismic data has shown it to be more consistent with a volcanic event¹³. Additionally, perturbations in crustal fluid pressure and temperature recorded at Ocean Drilling Program (ODP) sites on the eastern flank of the Endeavour segment indicated that the seismic energy generated by the earthquake swarm could account for only a small fraction of the estimated 12 cm displacement¹⁴. Thus, most of the displacement was aseismic, and occurred slowly enough not to produce seismic waves. This is consistent with slow magma movement.

CO₂ and He are magmatic volatiles that are strongly enriched in areas with active magma chambers, such as Axial volcano¹⁵, Loihi seamount¹⁶ and the eruptive area at 9° 50′ N on the East Pacific Rise^{17,18} (EPR). The concentrations of these two volatiles were relatively stable in fluids from five selected sulphide edifices (Hulk, Dante, Bastille, Puffer and Sully) in the MEF during the period 1991–98. However, in 1999 their concentrations dramatically increased and then declined to near pre-event values in 2000 (Fig. 1a, b; Supplementary Table 1). If we assume that the high-temperature reaction zone beneath Endeavour is about 2.5 km below the sea floor¹¹, the basaltic saturation value for CO₂ at this depth is 5 mmol kg^{−1} (ref. 19). Field data from the Juan de Fuca ridge indicate that CO₂ concentrations in most basalts range from saturation to twice saturation²⁰. On the basis of Li/Cl ratios, the water/rock ratio before and after the 1999 event remains relatively unchanged, and is about 2 (refs 3, 21; Fig. 1c; Supplementary Table 1). Assuming CO₂ values twice saturation, and allowing for the complete extraction of CO₂ in a single-pass system, rock extraction can only account for a fluid concentration of 5 mmol kg^{−1}. A tectonic event exposing new rock will not increase the CO₂ concentration without a reduction in the water/rock ratio. Water/rock ratios approaching 0.1 would be required to achieve the high CO₂ concentrations measured in 1999. As such low values were not seen, a magmatic source is required for the increased CO₂.

The significant drop in Cl concentration at Sully and Puffer immediately following the 1999 event (Fig. 1d) could be caused by increased phase separation induced by heat from either a tectonic or magmatic event. However, phase separation alone cannot explain the changes in gas concentration. Before 1999, these sulphide structures already exhibited low Cl concentrations owing to continuing phase separation. Increased phase separation should further increase the concentrations of gases, but this did not happen for all gases. For example, there was only a small increase in N₂ and Ar concentrations in 1999 beyond that seen in previous years (data not shown), whereas there was a 5-fold increase in CO₂ at Puffer in 1999 relative to 1998 (Fig. 1a). N₂ and Ar in vent fluids are primarily derived from circulating sea water and, except for a small Ar contribution from the mantle²², behave conservatively in the hydrothermal circulation cell. The relatively unchanged values for N₂ and

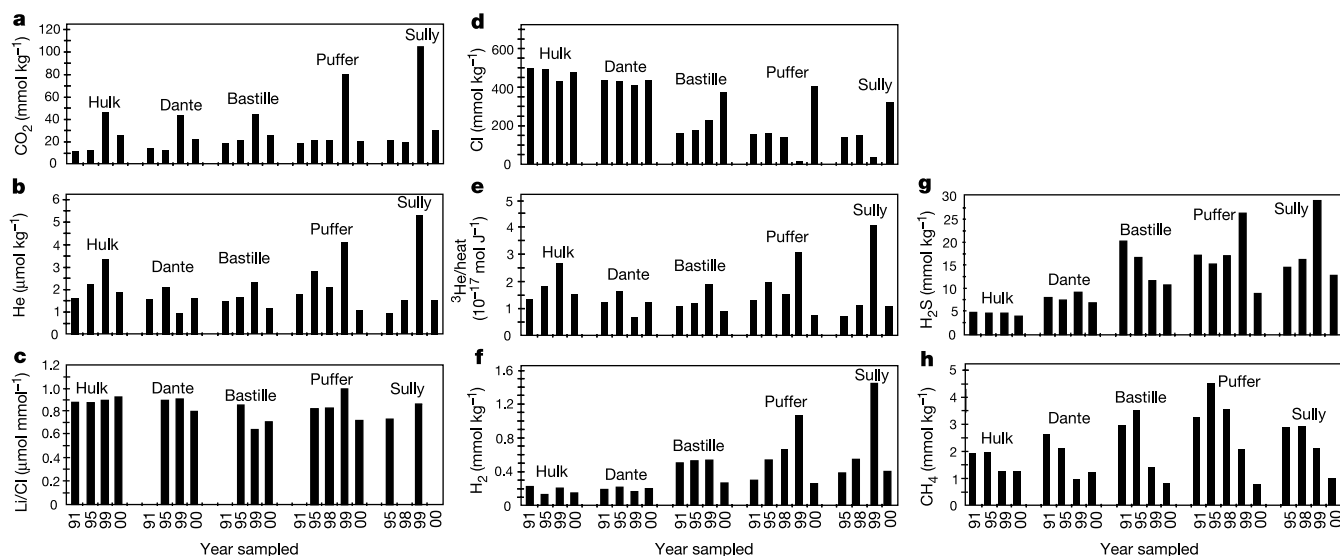


Figure 1 Time-series data for five of the major sulphide structures within the MEF. Shown are the concentrations of CO₂ (a), He (b), Cl (d), H₂ (f), H₂S (g) and CH₄ (h); also shown are the ratios Li/Cl (c) and ³He/heat (e). The values are ‘endmember’ concentrations. During sampling, small amounts of ambient sea water are entrained. This effect is removed by extrapolating to zero Mg concentration, as pure hydrothermal fluid contains

no Mg. The earthquake swarm started on 8 June 1999. The 1999 fluid samples were collected in September with DSV *Alvin* using gas-tight titanium samplers. The gases were extracted under vacuum at sea, and sealed in break-seal glass ampoules for analysis on shore by gas chromatography.

Ar indicate that the increased degree of phase separation in 1999 does not account for the large increases in CO_2 and He at Sully and Puffer (Fig. 1a, b). A shift to supercritical phase separation (brine condensation) at depth²³ could reduce the Cl content of the vapour phase without significantly altering the gas content. This, coupled with the earlier discussion of CO_2 extraction from basalt, indicates input from a magmatic source.

Before 1999, both Sully and Puffer had Cl concentrations ($135\text{--}160\text{ mmol kg}^{-1}$) much below that of sea water (540 mmol kg^{-1}). These low Cl values are indicative of a stable phase separation condition before the 1999 event. In 1999, however, the Cl concentration in MEF fluids decreased at four of the five sites, most dramatically at Puffer and Sully (Fig. 1d). In contrast, the Cl concentration at Bastille (near Puffer and Sully) increased in 1999, indicating the complexity of the effect of the event on the nature and degree of phase separation. In 2000, the Cl concentrations measured at Puffer, Sully and Bastille were more than double their respective concentrations seen before the event. Similar behaviour in Cl concentration after a magmatic event was seen at 'A' vent in 1991 on the EPR²⁴. Conversely, an apparent tectonic event on the EPR in 1995 had relatively little effect on Cl concentrations at nearby vents²⁵.

Without a decrease in water/rock ratio, elevated ^3He /heat ratios can also be taken to be a clear indicator of recent perturbation by magmatic activity. In all instances where data are available, the ^3He /heat ratio of fluids sampled after sea-floor volcanic events has been found to be elevated above the mean value of $0.5 \times 10^{-17}\text{ mol J}^{-1}$ for 'mature' hydrothermal fluids and to decline with time after the event²⁶. At MEF, the ^3He /heat ratio increased significantly at four of the five structures studied after the 1999 event, and then decreased back to, or below, the values recorded in preceding years (Fig. 1e). Our measurements at Endeavour were made approximately 4 months after the seismic event, and again at 1 year after the event. The relatively rapid decrease of CO_2 and He concentrations and ^3He /heat ratio back to pre-event levels within one year is similar to that observed at the 'A' vent at $9^\circ 46.5'\text{ N}$ on the EPR. At 'A' vent, the rapidly changing volatile concentrations (Fig. 2) were associated with the emplacement of a dyke about 200 m below the sea floor just two weeks before the collection of the first samples²². The rapid decrease in ^3He /heat ratio and corresponding increase in CO_2 concentrations indicates that CO_2 and He were differentially released during this magmatic event. These are the only volatile data documenting the early stages of such an event. By 1992, the 'A' vent ^3He /heat ratio had decreased to values similar to that at Endeavour. Concomitant initial decreases in Cl

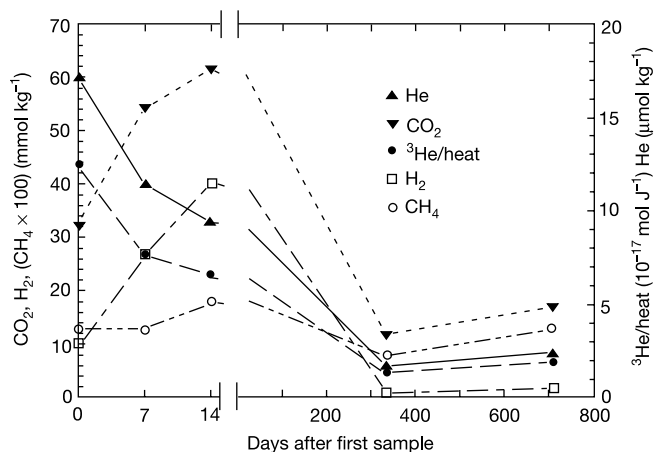


Figure 2 Time-series data for 'A' vent at 9°N on the EPR. Shown are concentrations of He, CO_2 , H_2 and CH_4 , and the $^3\text{He}/\text{heat}$ ratio. Samples were collected with DSV *Alvin* on 10, 17 and 24 April 1991; 10 March 1992; and 26 March 1994.

concentration and increases in H_2S were also noted at 'A' vent²².

In addition to an increase in CO_2 and He concentrations, the 1999 Endeavour event resulted in significant increases in H_2 and H_2S concentrations in some of the vents (Fig. 1f, g). Similar enrichments were noted at 'A' vent immediately following the dyke injection on the EPR, where H_2 values are the highest measured in hydrothermal fluids (Fig. 2). The concentration of H_2 in hydrothermal fluids is controlled by pressure, temperature, Cl concentration, and fluid–mineral equilibria²⁷. For a reaction zone at 400°C and 500 bar, a mineral assemblage consisting of anhydrite + anorthite, clinozomite, pyrite and magnetite would result in an H_2 concentration of about 0.2 mmol kg^{-1} (ref. 27). This fits well with the H_2 concentrations seen at Hulk and Dante (Fig. 1f). H_2 concentrations of the order of 0.5 mmol kg^{-1} were reported at Dante in 1999 at a higher-temperature vent than the source of our sample²⁸. On the basis of these mineral equilibria, the significant increases in H_2 and H_2S at Puffer and Sully following the 1999 event indicate a shift towards a more reducing pyrite–pyrrhotite–magnetite (PPM) redox buffer (Fig. 3). The redox shift at 'A' vent was more extreme, with H_2 and H_2S levels resting solidly within the pyrrhotite stability field. These more extreme values probably result because the heat source was only 200 m below the sea floor²², and also represent incipient rock alteration by seawater-derived hydrothermal fluid²⁷. These field data are consistent with the experimentally determined redox progression of basaltic alteration by hydrothermal fluid²⁷, but increased reaction-zone temperature and decreased pressure associated with magma movement would also substantially increase H_2 and H_2S concentrations in equilibrium with basaltic rock²⁷.

Not all gases within these systems behave similarly during a magmatic event. CH_4 concentrations decreased at Endeavour in 1999, and continued to decrease at three sites in 2000 (Fig. 1h), in marked contrast to CO_2 , He, H_2 and H_2S . Since 1984, anomalously high CH_4 concentrations (attributed to sedimentary involvement) have been in evidence at Endeavour²⁹. The decoupling of CH_4 concentrations from the other gases confirms that the Endeavour CH_4 source is not magmatic. CH_4 concentrations at 'A' vent, however, did show a systematic temporal trend as a result of the dyking event (Fig. 2).

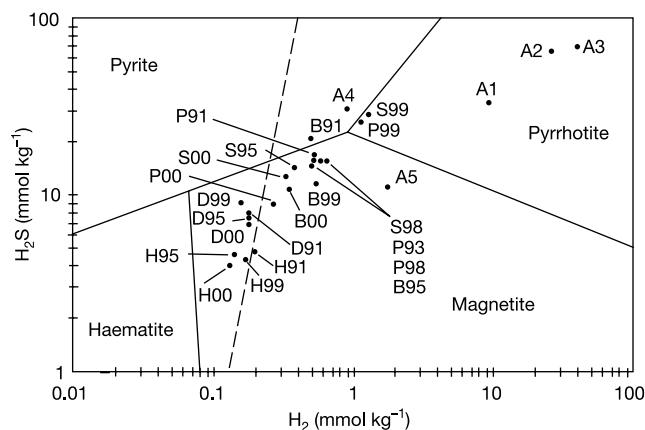


Figure 3 H_2 and H_2S concentrations plotted on the phase diagram for the mineral assemblage pyrite–pyrrhotite–magnetite (PPM) in the presence of anhydrite + anorthite–clinozomite (dashed line). Endeavour vents are designated by H (Hulk), D (Dante), B (Bastille), P (Puffer) and S (Sully) followed by the year sampled. The mineral stability fields represent the conditions at 400°C –500 bar, and are taken from ref. 27. These conditions are a good representation of the reaction zone at Endeavour, but the reaction zone at 'A' vent was much shallower (about 270 bar), which will shift the mineral stability fields to the right. As these low pressures promote large increases in H_2 and H_2S concentrations, the relationship of the 'A' vent samples to the stability fields is approximate.

Taking the data together—that is, the original seismic data¹³, the borehole data¹⁴, the large concentration increases in the magmatic gases CO₂ and He, and the increase in ³He/heat ratio—we conclude that the 8 June 1999 seismic swarm on Endeavour was the result of magma movement beneath the ridge crest, and not of a purely tectonic event. The data presented here confirm that magmatic events have profound effects on the characteristics of hydrothermal fluids. Magmatic activity causes transient fluxes of magmatic gases and shifts of pressure, temperature and redox conditions in the high-temperature reaction zone, thereby influencing fluid phase separation and causing large changes in the Cl concentration of hydrothermal fluids. Magmatic events can cause significant changes in the flux of ore-forming metals to the sea floor^{27,30}. The concentrations of CO₂, He and H₂ rise rapidly during magmatic events, but these concentration changes are not yet well constrained. The volatile flux during the first few months after magmatic events clearly needs to be better evaluated, as the resulting contribution of volatiles during this early period may rival that released by a mature hydrothermal system during an entire year. □

Received 2 December 2002; accepted 5 March 2003; doi:10.1038/nature01569.

- Kelley, D. S., Baross, J. A. & Delaney, J. R. Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* **30**, 385–491 (2002).
- Delaney, J. R., Robigou, V., McDuff, R. E. & Tivey, M. K. Geology of a vigorous hydrothermal system on the Endeavour segment, Juan de Fuca Ridge. *J. Geophys. Res.* **97**, 19663–19682 (1992).
- Butterfield, D. A. *et al.* Gradients in the composition of hydrothermal fluids from the Endeavour segment vent field: phase separation and brine loss. *J. Geophys. Res.* **99**, 9561–9583 (1994).
- Johnson, H. P. *et al.* Earthquake-induced changes in a hydrothermal system on the Juan de Fuca mid-ocean ridge. *Nature* **407**, 174–177 (2000).
- Kappel, E. S. & Ryan, W. B. F. Volcanic episodicity and a non-steady state rift valley along northeast Pacific spreading centers: evidence from Sea MARC I. *J. Geophys. Res.* **91**, 13925–13940 (1986).
- Rohr, K. M. M., Milkereit, B. & Yorath, C. J. Asymmetric deep crustal structure across the Juan de Fuca Ridge. *Geology* **16**, 533–537 (1988).
- White, D. J. & Clowes, R. M. Shallow crustal structure beneath the Juan de Fuca Ridge from 2-D seismic refraction tomography. *Geophys. J. Int.* **100**, 349–367 (1990).
- Wilcock, W. S. D., Archer, S. D. & Purdy, G. M. Microearthquakes on the Endeavour segment of the Juan de Fuca Ridge. *J. Geophys. Res.* **B 107**, doi:10.1029/2001JB000505 (2002).
- Lister, C. R. B. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Boström, K., Laubier, L. & Smith, K. L. Jr) 141–168 (Plenum, New York, 1983).
- Wilcock, W. S. D. & Delaney, J. R. Mid-ocean ridge sulfide deposits: evidence for heat extraction from magma chambers or cracking fronts? *Earth Planet. Sci. Lett.* **145**, 49–64 (1996).
- Detrick, R. S. *et al.* New multichannel seismic constraints on the crustal structure of the Endeavour Segment, Juan de Fuca Ridge: Evidence for a crustal magma chamber. *Eos* **83**, F1353 (2002).
- Lowell, R. P. & Germanovich, L. N. On the temporal evolution of high-temperature hydrothermal systems at ocean ridge crests. *J. Geophys. Res.* **99**, 565–575 (1994).
- Bohnenstiehl, D. R., Tolstoy, M., Dziak, R. P., Fox, C. G. & Smith, D. K. Aftershock sequences in the mid-ocean ridge environment: an analysis using hydroacoustic data. *Tectonophysics* **354**, 49–70 (2002).
- Davis, E. E., Wang, K., Thompson, R. E., Becker, K. & Cassidy, J. F. An episode of seafloor spreading and associated plate deformation inferred from crustal fluid pressure transients. *J. Geophys. Res.* **106**, 21953–21963 (2001).
- Butterfield, D. A., Massoth, G. J., McDuff, R. E., Lupton, J. E. & Lilley, M. D. Geochemistry of hydrothermal fluids from Axial Seamount hydrothermal emissions study vent field, Juan de Fuca Ridge: subseafloor boiling and subsequent fluid-rock interaction. *J. Geophys. Res.* **95**, 12895–12921 (1990).
- Sedwick, P. N., McMurtry, G. M. & MacDougall, J. D. Chemistry of hydrothermal solutions from Pele's Vents, Loihi seamount, Hawaii. *Geochim. Cosmochim. Acta* **56**, 3643–3667 (1992).
- Lupton, J. E., Lilley, M. D., Olson, E. J. & Von Damm, K. L. Gas chemistry of vent fluids from 9°–10° N on the East Pacific Rise. *Eos* **72**, F481 (1991).
- Lilley, M. D., Olson, E. J., McLaughlin, E. & Von Damm, K. L. Methane, hydrogen and carbon dioxide in vent fluids from the 9°N hydrothermal system. *Eos* **72**, F481 (1991).
- Dixon, J. E., Stolper, E. M. & Holloway, J. R. An experimental study of water and carbon dioxide solubilities in mid-ocean ridge basaltic liquids. Part I: Calibration and solubility models. *J. Petrol.* **36**, 1607–1631 (1995).
- Dixon, J. E., Stolper, E. & Delaney, J. R. Infrared spectroscopic measurements of CO₂ and H₂O in Juan de Fuca Ridge basaltic glasses. *Earth Planet. Sci. Lett.* **90**, 87–104 (1988).
- Kadko, D. & Butterfield, D. A. The relationship of hydrothermal fluid composition and crustal residence time to maturity of vent fields on the Juan de Fuca Ridge. *Geochim. Cosmochim. Acta* **62**, 1521–1533 (1998).
- Stuart, F. M. & Turner, G. Mantle-derived ⁴⁰Ar in mid-ocean ridge hydrothermal fluids: implications for the source of volatile and mantle degassing rates. *Chem. Geol.* **147**, 77–88 (1998).
- Seewald, J. S., Cruse, A. M. & Saccoccia, P. J. Aqueous volatiles in hydrothermal fluids from the Main Endeavour vent field: Temporal variability following earthquake activity. *Eos* **82**, F615 (2001).
- Von Damm, K. L. *et al.* Evolution of East Pacific Rise hydrothermal vent fluids following a volcanic eruption. *Nature* **375**, 47–50 (1995).
- Sohn, R. A., Fornari, D. J., Von Damm, K. L., Hildebrand, J. A. & Webb, S. C. Seismic and hydrothermal evidence for a cracking event on the East Pacific Rise crest at 9° 50' N. *Nature* **396**, 159–161 (1998).
- Lupton, J. E., Baker, E. T. & Massoth, G. J. Helium, heat, and the generation of hydrothermal event plumes at mid-ocean ridges. *Earth Planet. Sci. Lett.* **171**, 343–350 (1999).

- Seyfried, W. E. Jr & Ding, K. *Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions* (eds Humphris, S. E., Zierenberg, R. A., Mullineaux, L. S. & Thomson, R. E.) 248–272 (American Geophysical Union, Washington DC, 1995).
- Ding, K., Seyfried, W. E. Jr, Tivey, M. K. & Bradley, A. M. In situ measurement of dissolved H₂ and H₂S in high-temperature hydrothermal vent fluids at the Main Endeavour Field, Juan de Fuca Ridge. *Earth Planet. Sci. Lett.* **186**, 417–425 (2001).
- Lilley, M. D. *et al.* Anomalous CH₄ and NH₄⁺ concentrations at an unsedimented mid-ocean-ridge hydrothermal system. *Nature* **364**, 45–47 (1993).
- Von Damm, K. L. Chemistry of hydrothermal vent fluids from 9° – 10° N, East Pacific Rise: “Time zero,” the immediate post-eruptive period. *J. Geophys. Res.* **105**, 11203–11222 (2000).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank L. Evans and K. Roe for technical assistance, D. Kelley for a critical reading of the manuscript, W. Wilcock for discussions, and W. E. Seyfried Jr for the opportunity to join his 1999 cruise to Endeavour. The manuscript was improved by comments and suggestions from J. Seewald and R. Lowell. This work was supported by the National Science Foundation and in part by the National Oceanic and Atmospheric Administration VENTS Program.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.D.L. (e-mail: lilley@ocean.washington.edu).

A test of the unified neutral theory of biodiversity

Brian J. McGill

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

One of the fundamental questions of ecology is what controls biodiversity. Recent theory suggests that biodiversity is controlled predominantly by neutral drift of species abundances^{1–4}. This theory has generated considerable controversy^{5–12}, because it claims that many mechanisms that have long been studied by ecologists (such as niches) have little involvement in structuring communities. The theory predicts that the species abundance distribution within a community should follow a zero-sum multinomial distribution (ZSM), but this has not, so far, been rigorously tested. Specifically, it remains to be shown that the ZSM fits the data significantly better than reasonable null models. Here I test whether the ZSM fits several empirical data sets better than the lognormal distribution. It does not. Not only does the ZSM fail to fit empirical data better than the lognormal distribution 95% of the time, it also fails to fit empirical data better even a majority of the time. This means that there is no evidence that the ZSM predicts abundances better than the much more parsimonious null hypothesis.

The unified neutral theory of biodiversity (or UNTB, hereafter used to refer to either the theory or Hubbell's 2001 book¹) predicts that species abundance distributions (SADs) should follow the ZSM (see Box 1). To test whether the ZSM fits the data better than the null lognormal hypothesis, I used two data sets. First, I used the North American Breeding Bird Survey^{13–15} (BBS), for the simple reason that it is one of the few data sets that has replicates (in this case different sites) taken with identical sampling methodologies. Replicates are important for testing statistical significance. Since the ZSM has been applied primarily to intensively sampled data, I averaged the BBS data over a five-year period (1996–2000). Thus, rare species that show up only once in five years are included. I took 100 replicates by randomly selecting 100 routes that were rated to be of high quality for all five years. As an alternative data source, I also used the well-known Barro Colorado Island (BCI) tree data set, in

which trees are sampled in a 50-ha plot in Panama^{10,16}. This has the disadvantage of providing only a single data set, and hence not allowing an analysis of replicates. However, it is much larger (number of individuals and species) than a given BBS route, and

Box 1

Theory, testing and estimation of the ZSM

The SAD describes the relative abundances of different species within a community (see Fig. 3 for two ways to plot this information). The SAD nearly always shows many rare species and a few highly abundant species. Well over two dozen different probability distributions have been suggested as the ‘right’ SAD, with more appearing every year^{1,20–23}. The UNTB¹ has proposed that the SAD should follow a ZSM distribution.

The UNTB predicts that the species diversity and relative abundances of species can be explained by the neutral drift of the abundances of different species, in direct analogy to the neutral theory of molecular evolution²⁴. In particular, the UNTB suggests that the number of individuals in a community, J_M , is constant, and that at each time step, one random individual dies and is replaced either by a new species with probability ν , or by an offspring of one of the randomly selected remaining individuals with probability $1 - \nu$. Then the distribution of abundances (SAD) should be determined by the compound parameter $\theta = 2J_M\nu$, known as the fundamental biodiversity number. Hubbell calls this the ‘metacommunity’, and adds one further hierarchical level, the local community. In the local community, of population size J (much smaller than J_M , but to an unspecified extent), speciation does not occur, but when an individual dies it is replaced by a random sample from the metacommunity with probability m , or within the local community with probability $1 - m$, where m is the migration probability. Technically, this lower hierarchical level is a modified ZSM, but UNTB calls it just the ZSM, and I will follow this convention. Thus, the UNTB predicts that the SAD will follow the ZSM, which is parameterized by three numbers: θ , J and m .

The strongest test of this hypothesis to date is in Hubbell’s recent book (referred to as UNTB here)¹, which has provided a number of examples showing that the ZSM fits empirical data. However, the method by which goodness-of-fit of the ZSM to the data is measured is by visual examination (see, for example, Fig. 5.9 of the UNTB¹). No objective measure of goodness-of-fit is reported. Usually, science demands a stronger test, such as the data fitting well by some objective measure (for example, r^2) or preferably fitting better or even statistically significantly better than some null hypothesis²⁵. A reasonable null hypothesis for SADs is the lognormal distribution. Owing to arguments based on the central limit theorem²⁶, there are good statistical, non-biological reasons to expect the lognormal distribution to fit SADs²⁷. In this paper, I test whether the ZSM fits empirical data statistically significantly better than the lognormal.

The ZSM does not have an analytical form, so we can’t use likelihood ratios or other analytical methods of significance²⁸. For this reason I use a data set with replicates (the BBS) to allow significance testing. Moreover, fitting the ZSM distribution is quite complicated. I implement the method described in UNTB¹ (pages 289–294). Details over and above Hubbell’s original description are found in the Supplementary Information. Briefly, the method uses a generator function to generate one random sample from the ZSM. If we generate many (for example, 100) samples, and average across these samples, then we have a good estimate of the ZSM for a given set of parameters. We can evaluate the fit for different parameters using a generic likelihood function, and then search for the two parameters θ and m , which maximizes the observed likelihood. The algorithm was written in C and Matlab, and highly tuned for performance. Despite this, it takes half an hour of computation on an Athlon 500-MHz machine to fit the BCI data. A plot of the calculated likelihood surface can be found in the Supplementary Information. In view of the computational intensity of this calculation, we use a very simple hierarchical iterative method to search for the maximum likelihood, which is adequate for this problem because of the relatively smooth nature of the likelihood surface.

is exhaustively sampled. Most importantly, it has been used extensively in discussions on SADs and the ZSM^{1,2,10}.

The lognormal and ZSM distributions were fitted to the data using maximum-likelihood methods. Fitting the lognormal distribution to the data is quite simple; one merely log-transforms the data and then uses the usual methods for fitting the normal distribution (using the sample mean and a bias-corrected sample variance). Details of the estimating methods used for the ZSM can be found in Box 1 and the Supplementary Information. Source code is available from the author at <http://www.brianmcgill.org/zsmcode.html>. Eight measures of goodness-of-fit were then calculated and compared (described in the legends of Tables 1–3). Fits for the truncated lognormal were also calculated, but were found to perform slightly worse than the lognormal (although better than the ZSM in general), and are not presented. If the ZSM is better than the lognormal in 50% of cases (50 routes), then it performs better than the null. If it is better 95% of the time (95 out of 100 routes), then it is statistically significantly better.

Several issues were encountered while developing the fitting routines for the ZSM that have scientific interest and are described here: (1) Time to equilibrium. The ZSM fitting routine generates a ‘metacommunity’ population of very large size, takes a sample of the size of the local community, and then iterates the appropriate equations (see page 86 of ref. 1) for a number of time steps (individual deaths). No indication of what fixed number of time steps is appropriate has been published. My own experiments show that it can be large. For a local community of 1,600 individuals, it can take 10^7 time steps (individual deaths) to reach equilibrium. For a local community of 10,000 individuals, it can take 10^6 individual deaths (Fig. 1). Assuming a tree mortality rate of 1% per year (most studies find 1–2% per year under normal conditions^{17,18}), 10^6 time steps translates into 10,000 years just to reach equilibrium. Even

Table 1 Goodness-of-fit measures for BBS data

	r^2	r^2 MC	r^2 corr.	K
Lognormal	1.00 (0.99,1)	0.98 (0.93,0.99)	0.98 (0.96,1)	0.10 (0.06,0.15)
ZSM	0.99 (0.95,1)	0.89 (0.67,0.97)	0.97 (0.92,0.99)	0.26 (0.20,0.37)
ZSM beats lognormal	4%	4%	23%	0%

Compares goodness-of-fit for lognormal and ZSM distributions for 100 routes from the BBS. In each cell, the first number represents the mean across 100 routes; the numbers in parentheses represent the 2.5 and 97.5 percentiles (similar to a 95% confidence interval). The last row indicates the percentage of routes out of 100 where ZSM scores better than the lognormal. These four measures of goodness-of-fit are all based on calculating the observed cumulative distribution function (CDF), as is done in the Kolmogorov–Smirnov test. This is then compared with the theoretically predicted CDF using three measures of r^2 —raw r^2 , mean-corrected r^2 (r^2 MC) and square of the Pearson correlation r (r^2 corr.)—and the traditional Kolmogorov–Smirnov statistic, K , measuring greatest absolute vertical deviance.

Table 2 χ^2 goodness-of-fit measures for BBS data

	χ^2	χ^2 Preston log ₂ bins	χ^2 10 + 1 bins	χ^2 5 bins
Lognormal	19.7 (5.2,63.7)	8.97 (2.71,17.60)	11.1 (3.6,22.9)	8.64 (1.27,34.74)
ZSM	Inf. (15.9,NaN)	Inf. (12.912,Inf.)	19.0 (7.6,NaN)	Inf. (5.949,NaN)
ZSM ≠ INF	7	96	95	47
ZSM	20.7	24	19	12.9
ZSM beats lognormal	28.6%	0%	3.16%	6.38%

As in Table 1, but for various χ^2 measures of goodness-of-fit based on various bin schemes. The χ^2 statistic is undefined when the expected number of occurrences in a bin is zero. This causes the Inf. (infinite or ∞) and NaN (0/0) results. Although it biases results towards the ZSM, I also report the results for cases where Inf. and NaN results are thrown out. The number of remaining cases out of 100 is reported in the third row, and the average χ^2 value for just these cases is reported in the fourth row; the percentage of times the ZSM beats the lognormal is summarized in the last row. The four binning schemes, from left to right, are: (1) divide the observed range into 10 bins on an arithmetic scale; (2) divide the observed range into bins on a logarithmic scale according to Preston’s rules; (3) divide the lower 80% of the data into 10 arithmetic bins and include the upper 20% in one bin; and (4) similar to method 1 but with only 5 bins. Methods (2)–(4) are all attempts to ensure that the right-most bins are populated despite data that include very few species with very high abundances.

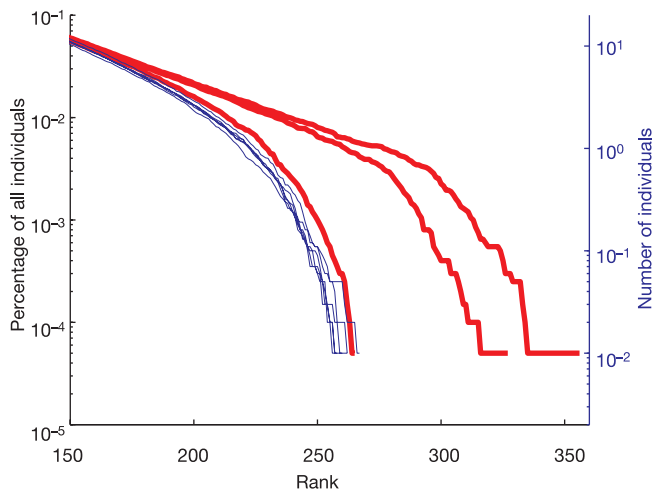


Figure 1 Approach to equilibrium of a local community. This is the lower right corner of a standard rank-abundance plot (see Fig. 3a for a full rank-abundance plot) with abundance on the vertical axis and rank on the horizontal axis (each point is one species; ranks 1–150 were omitted because they were identical for all nine lines; rank 1 is assigned to the most abundant species). The three red lines represent the state of the local community as it approaches equilibrium for 1×10^3 , 1×10^4 and 1×10^5 time steps (that is, individual deaths and going from right to left, respectively). We can see that by 1×10^5 deaths, the system is close to equilibrium, but is still noticeably different (at least over the range of rank 200–270). The right-most red line is closest to the source metacommunity (although the actual theoretical metacommunity would continue as a straight line rather than curving down). The six blue lines represent 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 and 1×10^7 time steps that have occurred in the local community since it became partially isolated from the metacommunity. The blue lines are all interpreted as being at equilibrium, as the 1×10^6 line is left-most, the 2×10^6 line is right-most, and the 1×10^7 line is in the centre. Within this noisy equilibrium, diversity drops down to about 250 species, jumps up to about 270 species for the next recorded time period (2×10^6 deaths), and oscillates in between for the rest of the time. This simulation had $\theta = 50$, $m = 0.1$ and $J = 20,000$ —close to the BCI data set.

assuming that a local community has 100,000 individuals, it would still take about 1,000 years. Thus, the time that the theory predicts it will take to reach local equilibrium may in fact be longer than the time over which environmental conditions are constant enough to support a given equilibrium. (2) Noisy equilibrium. The local community equilibrium is a very noisy one. Species diversities fluctuate over a range of about 10% of total species diversity (Fig. 1). This does not mean that the UNTB is not an equilibrium theory, but it is important to realize just how variable the equilibrium is. (3) Number of parameters and stopping rules. The methods used for generating the ZSM produce a large number of very rare species. In all of our runs, species were produced with abundances down to 0.01 individuals. This produced poor fits to the empirical data (too many rare species). The obvious solution is to truncate or stop the distribution somewhere. Looking at the figures in the UNTB, it is clear that the method used is to stop the distribution at the same number of species as the observed data. It would also have been reasonable to eliminate species with an abundance less than one. This (or any other stopping method) adds

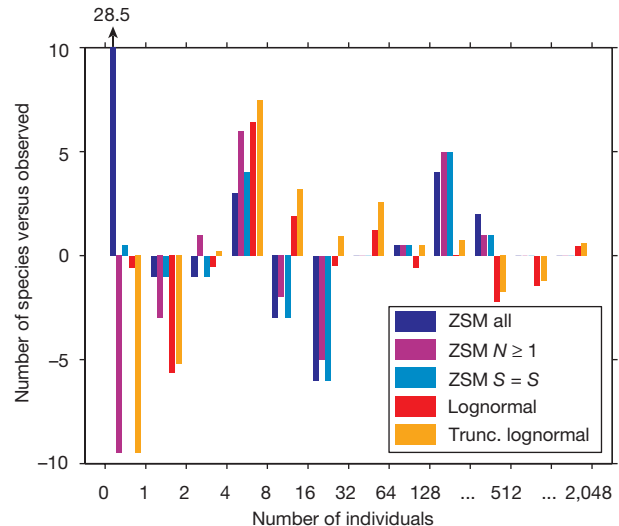


Figure 2 Deviations from observed abundances depends on the stopping rule. This figure shows deviations from observed abundances by Preston $\log_2$ bins for various stopping rules on the ZSM and lognormal. A positive value indicates predicting more species than observed. The vertical axis is truncated as the left-most bar should actually go up to 28.5. Trunc., truncated.

a third, undocumented fitting parameter, which makes the ZSM less parsimonious than it first appears to be.

This issue of the stopping method is pivotal in one of the main claims to success of the ZSM—predicting the abundances of rare species better than the lognormal (see figures 5.7 and 5.8, and the associated discussion, in the UNTB¹). The comparison shown in the two figures is based on having the ZSM stop at the observed number of species. If the ZSM is stopped where abundance is ≥ 1 , then it underpredicts the number of species. If the ZSM is not stopped, then it badly overpredicts the number of rare species. When the lognormal is stopped where abundance is ≥ 1 , it is called the truncated lognormal. As mentioned above, my own tests show that the truncated lognormal fits worse than the lognormal (presumably because the truncated lognormal is right-skewed, whereas exhaustively sampled abundance data is left-skewed¹⁹). In figures 5.7 and 5.8 of the UNTB¹, Hubbell compares the stopping rule for the ZSM that causes it to have the best possible fit with the stopping rule for the lognormal that causes the worst possible fit (that is, the truncated lognormal; see Fig. 2). When I compare the best ZSM stopping rule with the best lognormal stopping rule (Fig. 2), differences are small. Moreover, differences are also small, with the ZSM sometimes performing worse, if we make parallel comparisons of untruncated ZSM with untruncated lognormal, or of truncated ZSM with truncated lognormal. This sensitive dependency of the ZSM fit on the stopping rule means that we must examine carefully the statistical and scientific validity of stopping on the basis of the number of species. In the rest of this paper, I use the stopping rule giving the best fit in each case (having the same number of species as observed data for the ZSM as was done by Hubbell, and being untruncated for the lognormal).

The results of comparing goodness-of-fit for the BBS bird data

Table 3 Goodness-of-fit measures for the BCI tropical tree data

	r^2	r^2 MC	r^2 corr.	K	χ^2	Preston $\log \chi^2$	χ^2 10 + 1 bins	χ^2 5 bins
Lognormal	1	0.996	0.997	0.0565	12.7	5.45	10.3	4.57
ZSM	0.998	0.979	0.994	0.154	NaN	24.6	22.8	Inf.
ZSM beats lognormal	No	No	No	No	NA	No	No	No

Goodness-of-fit measures are as in the previous two tables. Only one data set is used, so no confidence intervals are presented. The last row indicates whether the ZSM scores better than the lognormal. NA, not applicable.

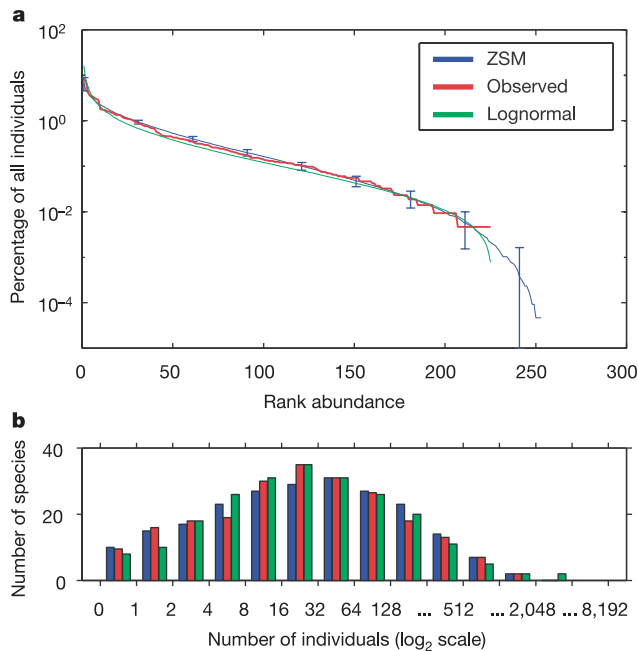


Figure 3 Comparing the fitted and actual BCI data. **a**, A standard rank-abundance plot (with abundance on the vertical axis and rank on the horizontal axis) of the estimated ZSM distribution versus the lognormal distribution versus actual data from the BCI data set. The error bars give one standard deviation above and below the fitted curve for the ZSM. **b**, A binned histogram using the same colour scheme as in **a** and with binning on a $\log_2$ scale. Thus, the bars between the tick marks for 2 and 4 indicate how many species with two to four individuals were calculated. Species with exactly two individuals are split between the 1–2 and 2–4 bins and so forth for all bin boundaries.

are summarized in Tables 1 and 2, and for the BCI tree data in Table 3. The ZSM performs reasonably well if the only criterion is to provide a good fit. However, when compared with the null hypothesis of the lognormal, the ZSM fares poorly. The average fits for the ZSM for all three types of r^2 are lower than the lognormal, and the ZSM beats the lognormal on only about 5–25% of the cases. It does much worse on the K (Kolmogorov–Smirnov) statistic, with an average K of 0.26 compared with 0.10 for the lognormal, and beats the lognormal in no case. Interpreting the χ^2 results is more complicated. The χ^2 statistic is undefined when the expected number of observations is zero. Because the ZSM distribution can be calculated only approximately, it regularly produces expected counts of zero. This results in either infinite χ^2 ($x/0$) or undefined χ^2 ($0/0$). Although such cases (especially the infinite) should probably count against the ZSM (they do not occur for the lognormal), I threw out all such cases and then compared the results with the lognormal for the remaining cases. Although this should bias the results towards the ZSM, the ZSM performs much worse than the lognormal on all χ^2 measures.

The 100 BBS routes averaged 77.6 species per route (with a 95% range of 32–105). These same routes averaged 3,875 individuals (over five years) with a 95% range of 1,325–7,023. The parameters for the ZSM on this data were as follows: the fundamental biodiversity number, θ , averaged 18.46 with a 95% range of 5.9–31.9, whereas migration probability, m , averaged 0.38 with a 95% range of 0.048–1. Interestingly, there was a strong correlation between species diversity, S , and the fundamental biodiversity number ($\theta = -3.94 + 0.289S$, $P < 0.0001$, $n = 100$, $r^2 = 0.48$; visual analysis of residuals indicates that untransformed linear regression is appropriate).

The results for the BCI tree data set are similar, although we cannot proceed to replicated comparisons against a null hypothesis,

because we have only one data set. The lognormal beats the ZSM on all measures of goodness-of-fit. Our estimated parameters were $\theta = 48.5$ and $m = 0.079$, resulting in an estimated curve that gave a good fit to the data (Fig. 3). My estimated parameters are close to the UNTB's estimates¹ of $\theta = 50$ and $m = 0.1$, and the difference is presumably due to fitting the BCI data for a year or a cutoff in minimum tree size that is different from that used by Hubbell (the number of individuals is very different from what Hubbell reports). Nonetheless, to make sure that this did not cause a poor fit, I also calculated goodness-of-fit using Hubbell's estimates of $\theta = 50$ and $m = 0.1$, and it did not change the results materially (usually only in the third significant digit).

A few cautions are in order when interpreting these results. First, it must be concluded that estimating the parameters for the ZSM is still an art and not an exact process, so there may be some room for improvement. However, given that I explored many approaches and always chose the one that made the ZSM perform best, I suspect that any further improvements will be small compared with the degree to which it falls short of the lognormal. Second, it should be noted that the UNTB also makes predictions regarding species area curves and phylogenies, which I have not addressed. Finally, it is important to note that the success of the lognormal does not mean that community structure is random; it simply means that community structure is a function of several multiplicative processes (the number of processes need not be large for the central limit theorem to produce a shape close to lognormal).

This paper has shown that the central empirical test of the UNTB (superior fit of the ZSM to observed abundance data; see chapter 9 of the UNTB¹) is in fact not true. The ZSM does indeed fit well. However, given the fact that the lognormal performs better, we cannot give the UNTB any special priority over the two dozen other theories that fit abundance data reasonably well. It is also important to note the extent of the failure to fit better than the lognormal. By using a data set with consistent methods (BBS) and many sites, we were able to quantify the relative performance; the lognormal outperforms the ZSM roughly 95–100% of the time for six of eight measures of goodness-of-fit, and roughly 75% of the time for the remaining two measures. Indeed, given the lower number of parameters and the greater parsimony, we may well not wish to favour any distribution derived from a complex theory until it is shown to perform better than the lognormal. Moreover, the UNTB starts with assumptions that are known to be wrong: the competitive equivalence of species^{6,8}. If the model is shown to have predictive power despite this, then we must question whether these competitive differences are important. But until there is a prediction of the UNTB that is strongly (that is, compared with a null hypothesis) and successfully tested, we have no evidence that these factors are unimportant, and much evidence that they are indeed significant. □

Received 18 October 2002; accepted 25 March 2003; doi:10.1038/nature01583.
Published online 13 April 2003.

- Hubbell, S. P. *A Unified Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
- Hubbell, S. P. Tree dispersion, abundance and diversity in a tropical dry forest. *Science* **203**, 1299–1309 (1979).
- Hubbell, S. P. & Foster, R. B. in *Community Ecology* (eds Diamond, J. & Case, T.) 314–328 (Harper & Row, Cambridge, Massachusetts, 1975).
- Bell, G. Neutral macroecology. *Science* **293**, 2413–2418 (2001).
- de Mazancourt, C. Consequences of community drift. *Science* **293**, 1772 (2001).
- Enquist, B. J., Sanderson, J. & Weiser, M. D. Modeling macroscopic patterns in ecology. *Science* **295**, 1835–1837 (2002).
- Levine, J. M. Species diversity and relative abundance in metacommunities. *Trends Ecol. Evol.* **17**, 99–100 (2002).
- Abrams, P. A. A world without competition. *Nature* **412**, 858–859 (2001).
- Clark, A. Macroecology comes of age. *Trends Ecol. Evol.* **17**, 352–353 (2002).
- Condit, R. *et al.* Beta-diversity in tropical forest trees. *Science* **295**, 666–669 (2002).
- Terborgh, J., Foster, R. B. & Nunez, P. V. Tropical tree communities: a test of the nonequilibrium hypothesis. *Ecology* **77**, 561–567 (1996).
- Yu, D. W., Terborgh, J. & Potts, M. D. Can high tree species richness be explained by Hubbell's null model? *Ecol. Lett.* **1**, 193–199 (1998).

13. Price, J., Droege, S. & Price, A. *The Summer Atlas of North American Birds* (Academic, San Diego, 1995).
14. Robbins, C. S., Bystrak, D. & Geissler, P. H. *The Breeding Bird Survey: Its First Fifteen Years, 1965–1979* (US Department of the Interior, Fish and Wildlife Service, Washington DC, 1986).
15. Sauer, J. R., Hines, J. E. & Fallon, J. *The North American Breeding Bird Survey, Results and Analysis 1966–2000 Version 2001.2* (USGS, Patuxent Wildlife Research Center, Laurel, Maryland, 2001); available at (<http://pwrctftr.er.usgs.gov/mp/bbs/DataFiles/>).
16. Pyke, C. R., Condit, R., Aguilar, S. & Lao, S. Floristic composition across a climatic gradient in a neotropical lowland forest. *J. Veg. Sci.* **12**, 553–566 (2001).
17. Burslem, D., Whitmore, T. & Brown, G. Short-term effects of cyclone impact and long-term recovery of tropical rain forest on Kolombangara, Solomon Islands. *J. Ecol.* **88**, 1063–1078 (2000).
18. Laurance, W., Ferreira, L., Rankin-De Merona, J. & Laurance, S. Rain forest fragmentation and the dynamics of Amazonian tree communities. *Ecology* **79**, 2032–2040 (1998).
19. Gaston, K. J. & Blackburn, T. M. *Pattern and Process in Macroecology* (Blackwell Science, Oxford, UK, 2000).
20. Tokeshi, M. Species abundance patterns and community structure. *Adv. Ecol. Res.* **24**, 111–186 (1993).
21. Pielou, E. C. *Mathematical Ecology* (John Wiley & Sons, New York, 1977).
22. Harte, J., Kinzig, A. P. & Green, J. Self-similarity in the distribution and abundance of species. *Science* **284**, 334–336 (1999).
23. Dewdney, A. K. A dynamical model of communities and a new species-abundance distribution. *Biol. Bull.* **198**, 152–165 (2000).
24. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, UK, 1983).
25. McGill, B. J. Strong and weak tests of macroecological theory. *Oikos* (in the press).
26. Grimmett, G. R. & Stirzaker, D. R. *Probability and Random Processes* (Clarendon, Oxford, UK, 1992).
27. May, R. M. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 81–120 (Belknap, Harvard Univ. Press, Cambridge, Massachusetts, 1975).
28. Hilborn, R. & Mangel, M. *The Ecological Detective: Confronting Models with Data* (eds Levin, S. A. & Horn, H. S.) (Princeton Univ. Press, Princeton, New Jersey, 1997).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements I thank the numerous volunteers who spent many hours collecting the BBS data, and the professionals at Patuxent Wildlife Research Center, who have put the data into an accessible, high-quality format, as well as those who collected and published the enormously valuable BCI data set. I also thank M. Rosenzweig, B. Enquist and the students in their laboratories for stimulating my thinking in this area. The National Science Foundation provided funding for this research. B. Walsh was a useful sounding board on statistics. Finally, I thank S. Marx, M. Weiser, D. Faulk and W. Turner for suggestions on an earlier draft of this paper, especially M. Weiser for his clarity on the importance of stopping rules, and S. Marx for greatly increasing the clarity of the paper.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (e-mail: mail@briannmcgill.org).

Ecological interference between fatal diseases

P. Rohani*, C. J. Green†, N. B. Mantilla-Beniers† & B. T. Grenfell†

* *Institute of Ecology, University of Georgia, Athens, Georgia 30602-2202, USA*
 † *Zoology Department, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*

An important issue in population biology is the dynamic interaction between pathogens. Interest has focused mainly on the indirect interaction of pathogen strains, mediated by cross immunity^{1–4}. However, a mechanism has recently been proposed for ‘ecological interference’ between pathogens through the removal of individuals from the susceptible pool after an acute infection. To explore this possibility, we have analysed and modelled historical measles and whooping cough records. Here we show that ecological interference is particularly strong when fatal infections permanently remove susceptibles. Disease interference has substantial dynamical consequences, making multi-annual outbreaks of different infections characteristically out of phase. So, when disease prevalence is high and is associated with significant mortality, it might be impossible to understand

epidemic patterns by studying pathogens in isolation. This new ecological null model has important consequences for understanding the multi-strain dynamics of pathogens such as dengue and echoviruses.

The possibility that epidemics of unrelated pathogens might interact has been raised in the classical epidemiological literature⁵, but has not been explained. Recently, a new mechanism has been proposed for negative ecological interference between pathogens through the temporary removal of susceptibles, arising from infection by a competing pathogen and the ensuing quarantine period⁶. Interference should be particularly apparent in the violent recurrent epidemics of strongly immunizing childhood infections such as measles and whooping cough. However, recent parallel records of the two infections in England and Wales show equivocal evidence for interference, partly because of the relatively low pathogenicity of the infections⁶. Here we test for interference in older data, collected when measles and whooping cough were significant killers⁵. Because a fatal infection involves the permanent removal of susceptibles, we would expect the imprint of interference to be particularly strong.

We begin by exploring the predictions of simple models, based on extensions of the classic one-disease seasonally forced SEIR (susceptible–exposed–infectious–removed) model^{7–9}, with two important biological refinements (see Methods). First, the model includes a convalescent class⁶, within which disease-induced deaths can occur; and second, we model the dynamics of two diseases simultaneously, categorizing hosts according to infection history relative to each disease. We are interested primarily in evaluating the dynamical impact of quarantine and disease-induced mortality on the community of pathogens. As with single-disease models, the dynamics of this system are determined largely by the recruitment rate of susceptibles (that is, the population birth rate^{7,8,10}; Fig. 1). Very low/high per-capita birth rates result in annual epidemics, with biennial dynamics observed for intermediate levels⁸. When there is a disease-related mortality rate, ρ , the window of biennial behaviour is progressively delayed, with the period-doubling bifurcation taking place at higher birth rates (Fig. 1a). This is because high mortality due to one infection in effect lowers the recruitment rate of susceptibles for the ‘competing’ disease. In general, the bifurcation structure of the model is dictated by the infection with the higher transmission rate—in this case, measles⁶. Whooping cough epidemics, which in isolation would be rigidly annual for all parameter combinations, now follow the same pattern as measles (Fig. 1b–d).

When epidemics are annual because of a low/high birth rate⁸, seasonal forcing causes strong positive correlation between infection outbreaks. However, given biennial epidemics, measles and whooping cough outbreaks are negatively correlated (out of phase)—much more so than if their dynamics were independent⁶. We use this negative correlation between disease dynamics as an indicator of potential interference in data.

We test for interference effects in case fatality reports for measles and whooping cough from Aberdeen (1883–1900) and from 15 European cities in the years before (1904–1914) and after (1922–1932) the First World War (Fig. 2). These data encompass large demographic heterogeneities, both spatial (between cities) and temporal (in the different eras). There were also systematic declines in the measles- and whooping-cough-induced death rates between periods before and after the war, which, along with the wide range of birth rates, provide an excellent opportunity to test the interference hypothesis in different regions of parameter space.

Children are typically affected by many more microparasitic infections than just measles and whooping cough (mumps, rubella and chickenpox within the childhood infections alone). However, we have focused on these two infections, because interference is likely to be most pronounced when diseases have very similar mean ages at infection, as dictated by their basic reproductive ratio, R_0 (ref. 7). Of potential childhood infections prevalent in the eras

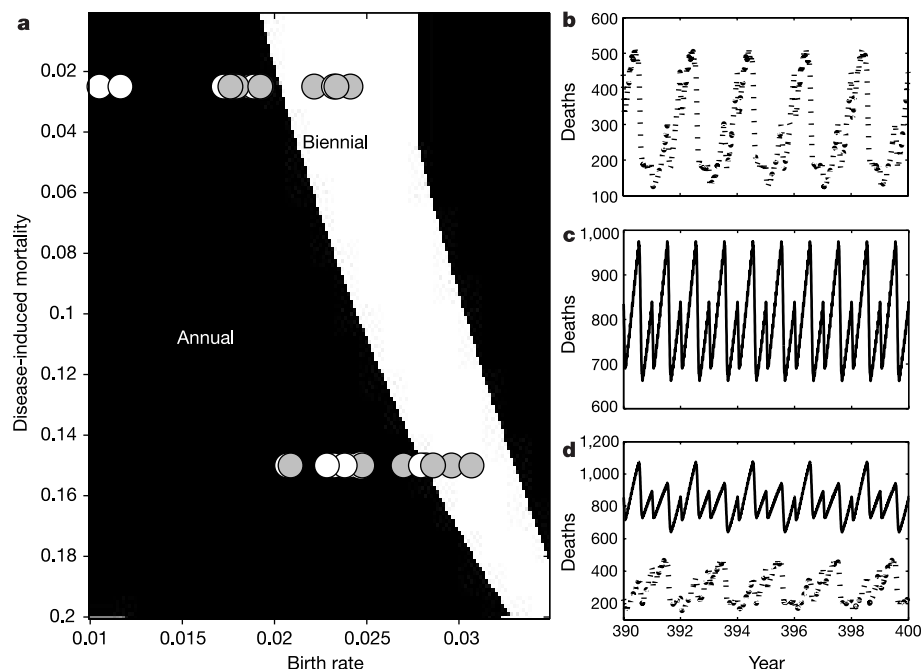


Figure 1 Analysis of the two-disease model dynamics. **a**, The bifurcation diagram illustrates the dynamics of the model as a function of the per-capita birth rate and the fraction of infecteds that suffer mortality. The black region highlights annual (and positively correlated) disease outbreaks; white regions identify biennial (negatively correlated) epidemics. The circles show the comparison with pre-war (mortality rate of 15%) and post-war (mortality rate of 2.5%) data. Negatively correlated data are

represented by grey circles; positive correlation is marked by a white circle. Panels **b–d** show time series for **b**, measles and **c**, whooping cough dynamics in single disease models compared with **d**, the two-disease model. The parameter values were: per-capita birth/death rate, 0.02; amplitude of seasonality, 0.3; incubation periods, 8 days; measles infectious period, 5 days; whooping cough infectious period, 14 days; measles convalescence period, 14 days; and whooping cough convalescence period, 21 days.

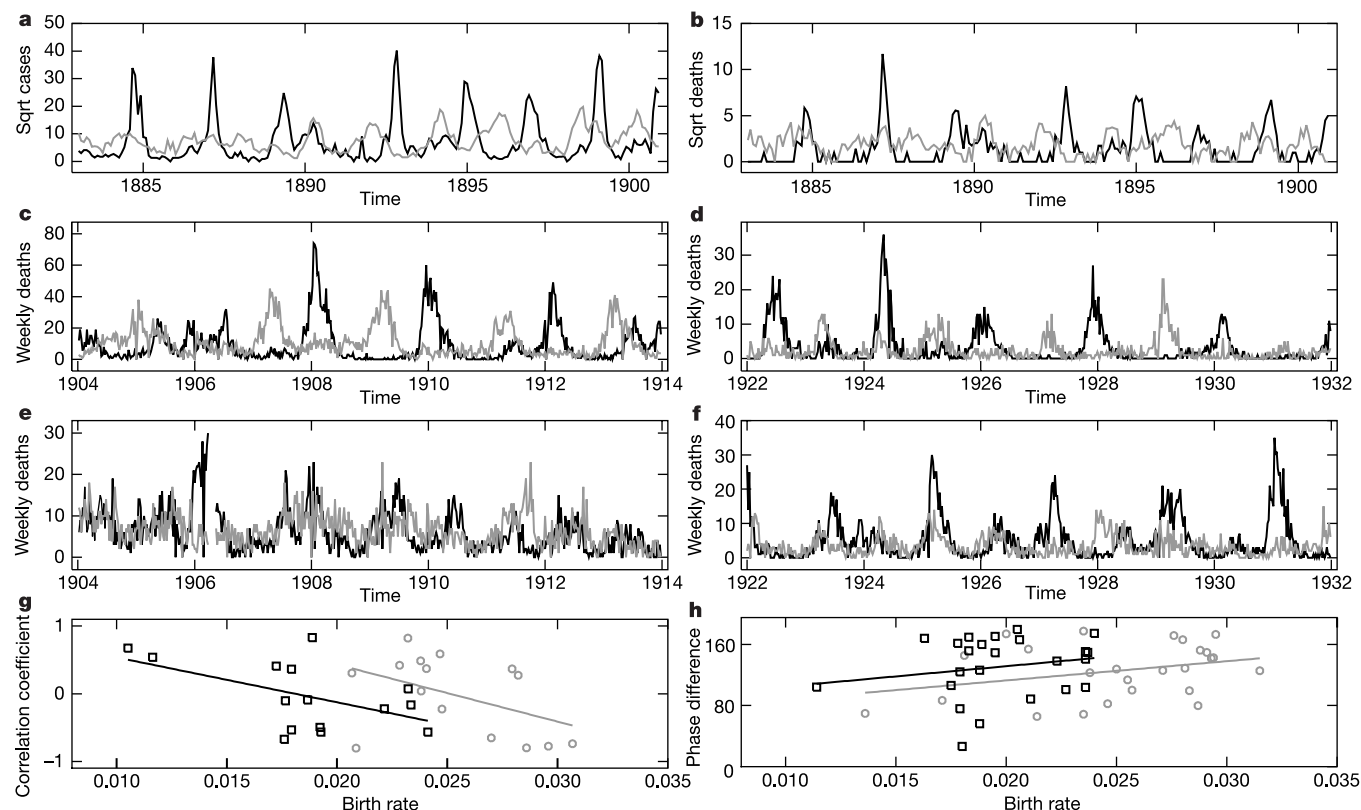


Figure 2 Weekly case fatality reports for measles (black) and whooping cough (grey) in five European cities. **a**, **b**, Case reports (**a**) and fatality (**b**) data for Aberdeen in the years 1882–1902. Sqrt, square root. **c–f**, Weekly deaths due to these two infections from 1904 to 1914 (**c**, **e**), and from 1922 to 1932 (**d**, **f**). Data are shown for Birmingham (**c**),

Glasgow (**d**), Berlin (**e**) and Liverpool (**f**). **g**, **h**, Correlation coefficients (left) between measles and whooping cough deaths, and their associated epidemic phase differences (right) as a function of the per-capita birth rates (grey symbols/lines refer to the 1904–1914 era; black symbols/lines represent the period 1922–1932).

studied, measles and whooping cough have similar mean ages at infection (approximately 5 years¹¹); this suggests a potential for interference, which would be strengthened by the 'cohort' effect of school entry¹². By contrast, measles and whooping cough would be expected to interact much less with the other childhood diseases present, which tend to infect children at later ages (see table 2 in ref. 11). The potential for interference between less infectious diseases (with a greater spread in the age at infection), or infections conferring imperfect immunity, is an interesting area for future work.

The time series show typical infection dynamics for measles with mainly biennial cycles^{11–14}, although on occasion annual and triennial dynamics are also observed. The whooping cough deaths over the same periods tend to exhibit either annual or biennial epidemics. Examples of the observed dynamics from five cities can be seen in Fig. 2. Figure 2a, b clearly shows that, in Aberdeen, both case reports and case fatalities of measles and whooping cough are biennial and out of phase. This pattern is also seen in Birmingham (pre-war; Fig. 2c) and Glasgow (post-war; Fig. 2d), whereas in pre-war Berlin and post-war Liverpool, they are predominantly annual and in phase (Fig. 2e, f). We summarize the relationship between measles and whooping cough epidemics by plotting the mean per-capita birth rates against the correlation coefficient and the mean phase difference, respectively (Fig. 2g, h; see Methods). For both measures of synchrony, low birth rates are generally associated with annual (positively correlated) epidemics, whereas higher birth rates generally give rise to biennial (negatively correlated) dynamics. These results accord qualitatively with the predictions of the interference model.

For a more detailed comparison, we plot the mean annual per-capita birth rate for each city in each era against the estimated case

fatality rate (conservatively estimated to be around 15% in the pre-war era, dropping to 2.5% after the First World War). We have shaded each symbol according to the correlation coefficient between measles and whooping cough epidemics; white (grey) indicates a positive (negative) correlation. As can be seen in Fig. 1, there is a good qualitative agreement between model predictions and the data. Despite the uncertainty about historical parameter values, we see that cities with negative inter-disease correlations generally fall within the biennial asynchronous regions predicted by the model, whereas the annual positively correlated cities tend to lie within the annual synchronized areas. These results, together with those presented in Fig. 2g, h, suggest a significant interference interaction between the infections.

The results of this study have potentially profound implications for the analysis of multiple strain dynamics¹⁵. Cross immunity is often crucial, but, especially in the case of fatal infections, may be superimposed on the 'ecological' effects outlined here. To illustrate the issues, we developed a three-strain (SIR) extension of the model and explored correlations between strains as a function of their respective convalescence periods and fatality probability (Fig. 3). The model findings highlight strain epidemics that become strikingly out of phase as either the convalescence period or fatality probability increases. Although our model is intended to be quite generic, its conclusions are consistent with the observed epidemics of the serotypes of dengue fever¹⁶ and the three strains of echoviruses¹⁵. In these systems, strains are found to fluctuate erratically (without a characteristic period) and out of phase with each other. Until now, dengue dynamics have been assumed to be due to either environmental/ecological determinants of the mosquito vector *Aedes aegypti*^{17,18}, or complex immunological interactions between the different strains—so-called antibody-dependent enhancement¹⁹. This simple null model clearly shows that analyses of immunologically or environmentally driven strain dynamics need to take into account the underlying ecological dynamics of infections¹⁵. The construction of models for combined strain and interference dynamics is clearly a priority for future work.

More broadly, we believe that interference is a striking example of non-stationary disease community dynamics tuned by exogenous forces. It gives us an outstanding opportunity to study competitive dynamics on the scale of individuals, and their large-scale consequences. It also underlines the potential pitfalls of studying the components of any ecological system in isolation. □

Methods

The model

We modified the classic one-disease SEIR model⁷. Initially, individuals are susceptible to both diseases. On contracting one disease, they enter the exposed class, and have negligible probability of simultaneously contracting the second disease (owing to increased non-specific immunological activity)²⁰. After the incubation period is over, the individual becomes infectious and still has negligible probability of contracting the other disease. Typically, when clinical symptoms appear, children enter a convalescent phase and are removed from school. Usually, this would be followed by complete recovery and lifelong immunity (although this may not always be the case for whooping cough^{21,22}). However, depending on age and condition (typically nutritional status), infection may be fatal owing to complications (such as pneumonia and encephalitis). Surviving individuals are then susceptible to the second disease only. It is only after experiencing both diseases that the individual enters the fully recovered class. A mathematical description of the model is provided in Supplementary Tables 1 and 2. Variables have double subscripts to denote their infection history—hence, for example, S_{SR} refers to those who are susceptible to infection 1 (measles) but have recovered from infection 2 (whooping cough).

To mimic the aggregation of children in schools, the model incorporates a seasonally varying contact rate—'term-time forcing'²³. For infection i , the contact parameter $\beta_i(t)$ depends on the time of year, being high on school days and low on other days. The parameters b_{0i} and b_1 are the disease-specific basic transmission rate and the seasonal amplitude, respectively.

$$\beta_i(t) = b_{0i}(1 \pm b_1)$$

European data

The data we present were published by the Registrar General's Weekly Returns. The correlation between case fatality data and infection levels is not well established. However, Butler presented data for London and Willesden in which there is a clear linear

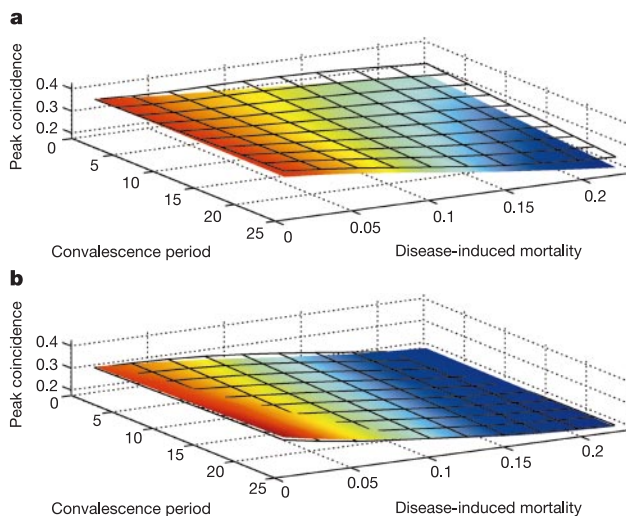


Figure 3 Temporal synchrony of strains, produced by a stochastic three-disease model. In **a**, $R_0 = 5$; in **b**, $R_0 = 10$. For any specific parameter combination, we estimated the pair-wise peak coincidence likelihood from a 250-year model time series, averaged them and fitted a spline (coloured surface). Peak coincidence, P , for a pair of time series is estimated by calculating the number of times they simultaneously peak (N) divided by the number of peaks, $\max[M_i, M_j]$, where M_i is the number of peaks for time-series i (ref. 30). To estimate confidence intervals, we randomized the phases of the time series 1,000 times and estimated P . The transparent surface shows a spline fitted to the 5th percentile of the bootstrapped values of P . As ρ (disease-induced mortality) or $1/\delta$ (convalescence period) increases, there is a decline in the peak coincidence levels, and strain epidemics become significantly out of phase (the coloured surface lies beneath the transparent one). Model parameters were per-capita birth/death rate, 0.02; amplitude of seasonality, 0.3; incubation plus infectious periods, 8, 9 and 10 days, respectively, for the three strains. Stochasticity was introduced into the model by making the transmission rate variable $\sim b_{0i} \times N(1, 0.05)$. This was estimated daily for each strain.

correspondence between case reports and fatality data. These data also establish that mortality rates are not affected by epidemic phase²⁴. Further confirmation of these results is provided by an analysis of the Aberdeen data (N.B.M.-B., P.R. and B.T.G., manuscript in preparation). Concerning infection-induced mortality rates, classic work by Butler²⁴, Bartlett²⁵, Creighton⁵ and others indicates significant mortality due to measles and whooping cough during these periods. Estimates of case fatality rates for measles vary widely, from 1–2% in the post-war era up to 46% pre-war^{14,26,27}, whereas estimates for whooping cough are in the 3–15% range²⁴.

Data analysis

These time series contain a substantial annual component and are further complicated by increasing population sizes over the two periods examined. Hence, analyses of the relationship between measles and whooping cough outbreaks were carried out on de-trended data. We used three separate methods. First, Pearson correlation coefficients were estimated for data aggregated over each epidemic year (October to October). Second, we carried out a linear regression of annual counts of measles against whooping cough and used the slope as a measure of synchrony. The results of this technique were qualitatively identical to those of the Pearson correlation, so we present only those. Finally, we also used Wavelet spectra to explore phase differences between filtered time series^{28,29}. Further information can be found in the Supplementary Information.

Bifurcation analysis

The bifurcation diagram was prepared by solving the system of ordinary differential equations (ODEs) for each parameter combination and characterizing the dynamics after a 190-year transient period was ignored. We used wrap-around initial conditions (whereby the final conditions for a run were used as the initial conditions for the next run). In the regions where the birth rates are relatively small, more than one attractor can coexist, although annual dynamics have the largest basin⁸. Note that in model dynamics, for some very restricted choice of initial conditions, it is possible to observe biennial dynamics that are in phase. However, work in progress has shown that infection dynamics become out of phase when stochasticity is introduced into the model, consistent with the findings of Kamo and Sasaki¹⁵.

Received 9 January; accepted 25 February 2003; doi:10.1038/nature01542.

- Gog, J. R. & Swinton, J. A. A status-based approach to multiple strain dynamics. *J. Math. Biol.* **44**, 169–184 (2002).
- Gupta, S., Ferguson, N. M. & Anderson, R. M. Chaos, persistence and evolution of strain structure in antigenically diverse infectious agents. *Science* **280**, 912–915 (1998).
- Gomes, M. G. M., Medley, G. F. & Nokes, D. J. On the determinants of population structure in antigenically diverse pathogens. *Proc. R. Soc. Lond. B* **269**, 227–233 (2002).
- Dietz, K. Epidemiologic interference of virus populations. *J. Math. Biol.* **8**, 291–300 (1979).
- Creighton, C. A *History of Epidemics in Britain* (Cambridge Univ. Press, Cambridge, 1894).
- Rohani, P., Earn, D. J. D., Finkenstadt, B. F. & Grenfell, B. T. Population dynamic interference among childhood diseases. *Proc. R. Soc. Lond. B* **265**, 2033–2041 (1998).
- Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford, 1991).
- Earn, D. J. D., Rohani, P., Bolker, B. M. & Grenfell, B. T. A simple model for complex dynamical transitions in epidemics. *Science* **287**, 667–670 (2000).
- Rand, D. A. & Wilson, H. B. Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics. *Proc. R. Soc. Lond. B* **246**, 179–184 (1991).
- McLean, A. & Anderson, R. Measles in developing countries part I. Epidemiological parameters and patterns. *Epidemiol. Infect.* **100**, 111–133 (1988).
- Anderson, R. M. & May, R. M. Directly transmitted infectious diseases: control by vaccination. *Science* **215**, 1053–1060 (1982).
- Schenzle, D. An age-structured model of pre- and post-vaccination measles transmission. *IMA J. Math. Appl. Med. Biol.* **1**, 169–191 (1984).
- Rohani, P., Earn, D. J. D. & Grenfell, B. T. Opposite patterns of synchrony in sympatric disease metapopulations. *Science* **286**, 968–971 (1999).
- Butler, W. Measles. *Proc. R. Soc. Med.* **6**, 120–153 (1913).
- Kamo, M. & Sasaki, A. The effects of cross-immunity and seasonal forcing in a multi-strain epidemic model. *Physica D* **165**, 228–241 (2002).
- Wenjie, W. Control of dengue/dengue haemorrhagic fever in china. *Dengue Bull.* **21** (<http://w3.whosea.org/DengueBulletin21/ch3f.htm>) (1997).
- Focks, D. A., Brenner, R. J., Hayes, J. & Daniels, E. Transmission thresholds for dengue in terms of *Aedes aegypti* pupae per person with discussion of their utility in source. *Am. J. Trop. Med. Hyg.* **62**, 11–18 (2000).
- Hales, S., de Wet, N., Maindonald, J. & Woodward, A. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. *Lancet* **360**, 830–834 (2002).
- Kurane, I., Mady, B. J. & Ennis, F. A. Antibody-dependent enhancement of dengue virus infection. *Rev. Med. Virol.* **1**, 211–222 (1991).
- Behrman, R. E. & Kliegman, R. M. *Nelson Essentials of Pediatrics* (Saunders, Philadelphia, 1998).
- Cherry, J. D. Pertussis in adults. *Ann. Intern. Med.* **128**, 64–66 (1998).
- Miller, E. & Gay, N. Epidemiological determinants of pertussis. *Dev. Biol. Stand.* **89**, 15–23 (1997).
- Keeling, M. J., Rohani, P. & Grenfell, B. T. Seasonally forced disease dynamics explored as switching between attractors. *Physica D* **148**, 317–335 (2001).
- Butler, W. Whooping cough and measles. *Proc. R. Soc. Med.* **40**, 384–398 (1947).
- Bartlett, M. S. Measles periodicity and community size. *J. R. Stat. Soc.* **1**, 48–59 (1957).
- Soper, H. E. The interpretation of periodicity in disease prevalence. *J. R. Stat. Soc.* **92**, 34–73 (1929).
- Linnert, L. A *statistical report on measles notifications in Manchester, 1917–1951*. (Department of Mathematical Statistics, Manchester, UK, 1954).
- Grenfell, B. T., Bjornstad, O. N. & Kappey, J. Travelling waves and spatial hierarchies in measles epidemics. *Nature* **414**, 716–723 (2001).
- Torrence, C. & Compo, G. P. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **79**, 61–78 (1998).

30. Buonaccorsi, J. P., Elkington, J. S., Evans, S. R. & Liebholt, A. M. Measuring and testing for spatial synchrony. *Ecology* **82**, 1668–1679 (2001).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank O. Bjornstad, M. Boots, D. Gubler and H. Wearing for comments on this manuscript.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.R. (e-mail: rohani@uga.edu).

Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing

Shuang-Xi Ren^{*†‡}, Gang Fu^{*§‡}, Xiu-Gao Jiang^{||‡}, Rong Zeng^{†‡}, You-Gang Miao[†], Hai Xu[†], Yi-Xuan Zhang[†], Hui Xiong[§], Gang Lu^{*}, Ling-Feng Lu^{*}, Hong-Quan Jiang^{*§}, Jia Jia^{*}, Yue-Feng Tu^{*}, Ju-Xing Jiang^{||}, Wen-Yi Gu^{*}, Yue-Qing Zhang^{**}, Zhen Cai^{*}, Hai-Hui Sheng^{*}, Hai-Feng Yin^{*}, Yi Zhang^{*}, Gen-Feng Zhu^{*}, Ma Wan^{||}, Hong-Lei Huang^{||}, Zhen Qian^{*}, Sheng-Yue Wang^{*}, Wei Ma[†], Zhi-Jian Yao^{||}, Yan Shen^{||}, Bo-Qin Qiang^{||}, Qi-Chang Xia[†], Xiao-Kui Guo[§], Antoine Danchin[☆], Isabelle Saint Girons^{**}, Ronald L. Somerville^{††}, Yu-Mei Wen[#], Man-Hua Shi^{||‡‡}, Zhu Chen^{*§}, Jian-Guo Xu^{||} & Guo-Ping Zhao^{*†}

^{*} Chinese National Human Genome Center at Shanghai (CHGCS), 250 Bi Bo Road, Zhang Jiang High Tech Park, Shanghai 201203, China

[†] Bioinformation Center/Institute of Biochemistry and Cell Biology/Institute of Plant Physiology and Ecology/Research Center of Biotechnology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China

[§] Rui Jin Hospital/Department of Microbiology and Parasitology, Shanghai Second Medical University, 280 Chongqingnan Road, Shanghai 200025, China
^{||} National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention (ICDC, China CDC), P.O. Box 5, Changping, Beijing 102206, China

^{||} Chinese National Human Genome Center, Beijing, 707 North Yongchang Road, Yi Zhuang High Tech Park, Beijing 100170, China

[#] Department of Molecular Virology, Medical Center, Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032, China

[☆] HKU-Pasteur Research Centre, 8, Sassoon Road, Hong Kong, China

^{**} Unité de Bactériologie Moléculaire et Médicale, Institut Pasteur, 25, rue du docteur Roux, 75724 Paris Cedex15, France

^{††} Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, USA

[‡] These authors contributed equally to this work

^{‡‡} Deceased

Leptospirosis is a widely spread disease of global concern. Infection causes flu-like episodes with frequent severe renal and hepatic damage, such as haemorrhage and jaundice. In more severe cases, massive pulmonary haemorrhages, including fatal sudden haemoptysis, can occur¹. Here we report the complete genomic sequence of a representative virulent serovar type strain (Lai)² of *Leptospira interrogans* serogroup Icterohaemorrhagiae consisting of a 4.33-megabase large chromosome and a 359-kilobase small chromosome, with a total of 4,768 predicted genes. In terms of the genetic determinants of physiological characteristics, the facultatively parasitic *L. interrogans* differs

extensively from two other strictly parasitic pathogenic spirochaetes, *Treponema pallidum*³ and *Borrelia burgdorferi*⁴, although similarities exist in the genes that govern their unique morphological features. A comprehensive analysis of the *L. interrogans* genes for chemotaxis/motility and lipopolysaccharide synthesis provides a basis for in-depth studies of virulence and pathogenesis. The discovery of a series of genes possibly related to adhesion, invasion and the haematological changes that characterize leptospirosis has provided clues about how an environmental organism might evolve into an important human pathogen.

Spirochaetes, morphologically unique in their coiled, slender and flexuous shape and related form of motility, form a major phylogenetic lineage (phylum) of eubacteria. *Leptospira*, an obligately aerobic, tightly coiled spirochaete, is the only genus other than *Borrelia*, *Treponema* and *Brachyspira* that is able to cause significant infection in mammals. The leptospires are physiologically chemoheterotrophic. They include the saprophytic *L. biflexa* and the pathogenic *L. interrogans*. The latter is known worldwide to be responsible for the water-borne zoonosis leptospirosis. Although antibiotic therapy is effective against the disease, it remains a serious threat in tropical and subtropical countries as well as in those cities where sanitation is substandard and where wild rats can serve as reservoirs when sewage disposal is poor¹.

Molecular and cellular studies on leptospires⁵ have focused on their dynamics of motility, biosynthesis of amino acids and lipopolysaccharide (LPS), outer-membrane proteins and other potential virulence factors. In contrast to *L. biflexa*, little in the way of genetic analysis has been reported for *L. interrogans*, owing to their fastidious cultivation requirements and the lack of genetic systems⁵. Previously, the genomic sequences of two pathogenic spirochaetes—*T. pallidum*, responsible for syphilis³, and *B. burgdorferi*, responsible for Lyme disease⁴—have been determined. We employed the whole-genome random sequencing method^{3,4,6} to sequence and analyse the genomic DNA of a representative virulent serovar type strain (Lai)² of *L. interrogans* serogroup Icterohaemorrhagiae (see Methods).

The *L. interrogans* genome (4,691,184 base pairs (bp); Fig. 1, Table 1) is much larger than either of the other two spirochaetes (1,138,006 bp for *T. pallidum* and 1,519,857 bp for *B. burgdorferi*, including plasmids). It consists of two circular chromosomes, a large one of 4,332,241 bp (CI) and a small one of 358,943 bp (CII), in good agreement with previous estimates⁵. More than 30 copies of repetitive DNA elements, including members of the IS1500 and IS1501 families, were distributed throughout the genome but few phage-related sequences were identified.

Both GC nucleotide skew ((G – C)/(G + C)) analysis and comparisons with the *ori* sequences of other bacteria were employed to

locate the replication origin of CI, whereas only GC nucleotide skew analysis was used to identify a putative replication origin on CII, as with *Vibrio cholerae*⁶. DnaA boxes were identified on the anti-clockwise side of *oris* for both CI and CII. In addition, *parAB* operons were identified on each side of the putative replication origins of both chromosomes (Supplementary Information 1).

In all, 4,768 putative genes were predicted, among them 37 genes for transfer RNAs (Supplementary Information 2-1). Previous reports⁵ indicated that in strains Ictero No. 1, Verdun and RZ11 of *L. interrogans*, there were two sets each of genes encoding 16S ribosomal RNA (*rrs*) and 23S rRNA (*rrl*). However, besides the two *rrs* genes, we identified only one gene each encoding 5S (*rrf*) and 23S rRNAs respectively. The extraordinarily low number of tRNA and rRNA genes might well account for the fastidious growth of *L. interrogans*.

Among the 4,727 protein-coding sequences (CDSs), 4,360 lie on CI and 367 lie on CII, whereas all of the rRNA and tRNA genes were found on CI (Table 1). Although most of the genes required for growth and viability are located on CI, some essential genes lie on CII. Besides the previously recognized *metF*⁷ (LB002) and *asd*⁵ (LB355), it is significant to recognize an *ndh* gene (LB036), encoding NADH dehydrogenase, and clusters of genes involved in a nearly complete pathway for the *de novo* biosynthesis of haem. These data, therefore, tend to support the view that CII is an authentic part of the genome that did not originate by lateral transfer.

On the basis of amino acid sequence similarity searches and/or domain analysis, biological functions have been assigned to about 44% of the CDSs (2,060), whereas 15% of the CDSs (715) either encode proteins of unknown function or are similar to unassigned CDSs predicted in other organisms. A total of 1,952 predicted CDSs (41%) failed to exhibit obvious similarity to any protein-coding genes of other organisms (Table 1). In particular, only 315 orthologues were shared by *L. interrogans*, *T. pallidum* and *B. burgdorferi* (Supplementary Information 3).

Some of the previously identified metabolic characteristics of leptospires, such as the absence of hexokinase¹, were confirmed by genomic analysis. A complete set of genes for a system of long-chain fatty-acid utilization, a tricarboxylic acid cycle and a respiratory electron transport chain were identified in *L. interrogans*; this was consistent with the notion that the organism generates ATP by oxidative phosphorylation (Fig. 2). In contrast, none of the aforementioned genes are present in *T. pallidum* or *B. burgdorferi*, in which ATP can be generated only by sugar fermentation by means of the Embden–Meyerhof pathway⁵. Because *L. interrogans* cannot utilize sugars as carbon sources, anaplerotic reactions are essential for gluconeogenesis. We failed to identify genes encoding glucose-6-phosphate dehydrogenase, one of the key enzymes of the phosphogluconate pathway. Neither of the two key enzymes of the glyoxylate pathway, isocitrate lyase and malate synthase, were present, although these two enzymes were detected in *L. biflexa*¹. However, we did identify all the genes encoding enzymes for gluconeogenesis from glycerol (Fig. 2), including phosphoglucose isomerase, as previously reported¹. In addition, genes encoding enzymes likely to be involved in the oxidative carboxylation of acetyl-CoA to succinyl-CoA through the 3-hydroxypropionate pathway⁸ were recognized (Fig. 2). Intermediates of carbohydrate metabolism are therefore likely to be synthesized by means of the tricarboxylic acid cycle and the non-oxidative pentose phosphate pathway (Fig. 2). Genes encoding transhydrogenase (*pntA* and *pntB*) were identified. These enzymes could catalyse the formation of sufficient NADPH for anabolic processes at the cost of protonmotive force generated by an NADH dehydrogenase complex (Fig. 2). In this connection, one should emphasize that glycerol, together with the long-chain fatty acids, is present in EMJH medium (Johnson and Harris modification of the Ellinghausen and McCullough medium)¹ for better growth of *L. interrogans*.

In contrast to *B. burgdorferi* and *T. pallidum*, *L. interrogans*

Table 1 General features of the *L. interrogans* chromosomes

Features of <i>L. interrogans</i> chromosomes	CI	CII	CI + CII
Genome size (bp)	4,332,241	358,943	4,691,184
G + C (%)	36.00	36.10	36.00
Protein coding (%)	78.30	79.80	78.40
Protein-coding genes (no.)	4,360	367	4,727
CDSs with functional assignment	1,901	159	2,060
CDSs without functional assignment	2,459	208	2,667
Unknown functional proteins	140	6	146
Similar to unassigned CDSs predicted in other organisms	509	60	569
No significant similarity to CDSs predicted in other organisms	1,810	142	1,952
Gene density (bp per gene)	993	978	992
Average gene length (bp)	778	781	778
Average unknown gene length (bp)	557	567	558
Insertion sequence (no.)	18	12	30
IS 1500 family	7	1	8
IS 1501 family	1	4	5
Others	10	7	17
Ribosomal RNAs	4	0	4
Transfer RNAs	37	0	37

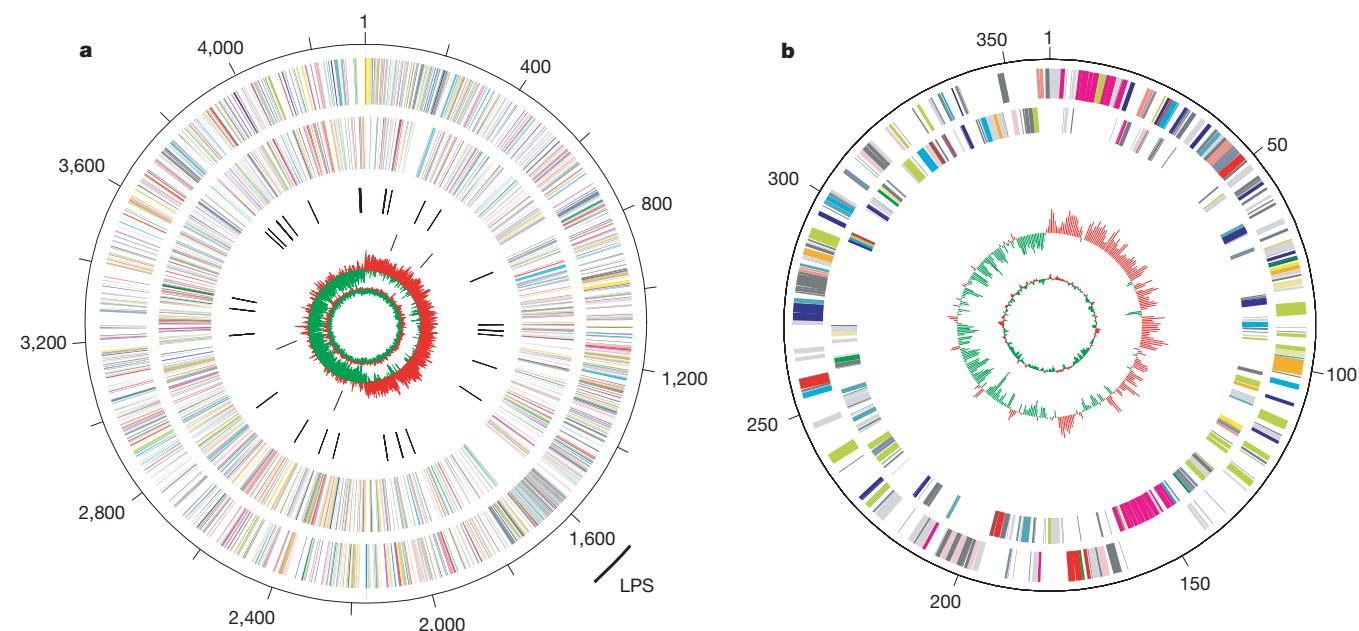


Figure 1 Circular representation of the *L. interrogans* strain Lai genome, with predicted CDSs. **a**, Large chromosome (CI); **b**, small chromosome (CII). The outer scale is shown in kilobases. Circles range from 1 (outer circle) to 6 (inner circle) for CI and from I (outer circle) to IV (inner circle) for CII. Circles 1/I and 2/II, genes on forward and reverse strand; circles 3, tRNA genes; circle 4, rRNA genes; circle 5/III, GC bias ((G-C)/(G + C)); red indicates values > 0; green indicates values < 0; circles 6/IV, G + C content. All genes are colour-coded according to functions: orange for amino acid biosynthesis, green for purines, pyrimidines, nucleosides and nucleotides, blue for fatty acid and phospholipid

metabolism, magenta for biosynthesis of cofactors, prosthetic groups and carriers, khaki for central intermediary metabolism, cyan for energy metabolism, orchid for transport and binding proteins, yellow for DNA metabolism, dark green for transcription, brown for protein synthesis, red for protein fate, green-yellow for regulatory functions, pink for cell envelope, salmon for cellular processes, navy for other categories, light grey for conserved, dim grey for hypothetical, slate grey for unknown function protein, and black for tRNA and rRNA.

encodes complete metabolic systems for amino acid and nucleotide biosynthesis, which is in agreement with previous work¹. Methionine biosynthesis in leptospire is similar to that in yeast¹, whereas the final step seems to be catalysed by a B₁₂-dependent homocysteine-N⁵-methyltetrahydrofolate transmethyrase, encoded by *metH*, rather than by a cobalamin-independent methionine synthase encoded by *metE* (Fig. 2). In this connection, the absence of several genes of B₁₂ biosynthesis from the *L. interrogans* genome accounts for the fact that this compound is an essential component of the EMJH semi-synthetic medium¹. It was proposed that a pyruvate pathway might be used by leptospire for isoleucine biosynthesis, either alone or together with the conventional threonine deaminase pathway¹. Because we failed to identify a gene encoding threonine deaminase but did find three putative *leuA* genes, we experimentally determined the substrate specificity of these enzymes (see Methods). The enzyme encoded by LA2202 is an isopropylmalate synthase (*leuA1*), whereas LA2350 encodes citramalate synthase (*cimA*). Although the enzyme encoded by LA0469 has some citramalate synthase activity, it is primarily an isopropylmalate synthase (*leuA2*).

The genomic information enhances our understanding of the mechanisms of virulence and pathogenesis in leptospirosis. As with most other pathogenic bacteria, *L. interrogans* possesses several genes related to the attachment and invasion of eukaryotic cells (*mce*, *invA*, *atsE* and *mviN*; Supplementary Information 2-2). The unique cellular shape and motility apparatus of spirochaetes provide these organisms with an additional method of achieving effective infection^{5,9}. We found at least 50 genes (not including chemotaxis genes) related to motility, accounting for more than 1% of the deduced CDSs (Fig. 2). Like *B. burgdorferi* and *T. pallidum*, *L. interrogans* uses FlaA sheath protein and FlaB core protein as the essential components of its endoflagellar filament⁷. Other bacteria^{5,10} employ FliC for this purpose. *L. interrogans* also has a

complete set of genes (Supplementary Information 2-2) for shape determination. In contrast to *B. burgdorferi*¹¹, the finely coiled spiral shape of leptospire is likely to be mainly attributable to the murein layer rather than the flagella¹².

Chemotaxis is generally acknowledged to be an important virulence factor for pathogenic bacteria. The chemotaxis system of *L. interrogans* (Fig. 2) is more complex than that of either *T. pallidum* or *B. burgdorferi*. The recognition of many genes (12 CDSs) encoding methyl-accepting chemotaxis proteins (MCPs) presumably reflects the extremely diverse environmental situations that a facultatively parasitic zoonotic bacterium can encounter. Employing secondary-structure prediction methods, 5 of the 15 CDSs with clear CheY-like response domains were designated *cheY* genes (Supplementary Information 4-1). However, only one such gene was located in a putative chemotaxis operon (*cheWABY*, LA1250-1253).

Leptospirosis virulence has been attributed in part to the effect of the leptospiral LPS¹. The nucleotide sequence of the locus encoding a set of enzymes for the biosynthesis of the O-antigen component of *Leptospira* LPS (*rfb* locus) is known for four serovars of two species⁵. We identified an *rfb* locus of 103 kilobases (kb) (Supplementary Information 5) in *L. interrogans* serovar lai. In agreement with findings in other *rfb* loci of leptospire, almost all of the 97 CDSs (LA1576 to LA1672), except three short ones, are encoded on the same strand (forward). About 30 kb of nucleotide sequence located at the 3'-proximal end of the locus is almost identical to its counterpart in serovar copenhageni (GB: U61226). Unlike *L. borgpetersenii* serovar hardjo, subtype hardjobovis, no IS elements were found within or flanking the *rfb* locus. We tentatively assigned a series of genes encoding O-antigen-processing enzymes within and outside the *rfb* locus by comparisons of predicted transmembrane patterns with genes characterized in other Gram-negative bacteria (Supplementary Information 4-2). This is a strong

haemostasis, so far found only in *Leptospira*, seems to specify an orthologue of paraoxonase (LA0399, *pon*). This protein might hydrolyse PAF through its arylesterase activity¹⁶. Because a *cola*¹⁷ gene (LA0872) encoding microbial collagenase has been identified, it is reasonable to propose that collagenase-mediated injury to the vascular epithelium during infection and the subsequent combined effects of the Vwa, PafAH and Pon proteins could lead to a loss of haemostasis, in addition to the proposed effects of LPS^{1,13}. This model is consistent with the clinical manifestations of leptospirosis, namely damage to the endothelial cell membranes of small blood vessels¹. It also might explain the observed sequelae of severe infections by serovar lai, such as massive pulmonary haemorrhage and fatal sudden haemoptysis¹.

Among eubacteria, spirochaetes are evolutionarily primitive^{9,18}. However, the fact that leptospires can survive either as saprophytes or as facultative parasites has presumably afforded them significant growth opportunities, although not without pressure for co-evolution in response to their environment or hosts. A BLAST analysis was performed to compare the best-hit distribution of protein homologues in representative eubacteria with the predicted proteomes of bacteria, virus (phage), archaea and eukarya. The result (Fig. 3) suggests that the genome of *L. interrogans* surpasses those of other bacteria in terms of the number of proteins with structural similarity to eukaryal and archaeal proteins that it encodes. In this respect, *L. interrogans* resembles *B. burgdorferi* and *Mycoplasma genitalium*. This raises several important evolutionary questions, including the possibility that lateral gene transfer, operating in parallel with standard gene evolution events, contributed to the emergence of an important human pathogen from an environmental bacterium. □

Methods

Source and culturing of study organism

The *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar lai type strain 56601 used in this study is maintained by the National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention (ICDC, China CDC), Beijing, China². For sequencing purposes, a single colony was picked from EMJH¹ soft agar and cultured in the same medium. The culture thus obtained was then subjected to morphological, serological, genetic and virulence analysis. The properties of the strain were in accordance with those of pathogenic *Leptospira*. For functional analysis, growth curves for *L. interrogans* in EMJH or Korthof¹ medium were measured turbidimetrically, and viable bacterial counts were determined by dark-field microscopy. Culture conditions were then developed to ensure that only mid-exponential-phase bacterial cultures were used for further experimentation.

Genome sequencing and analysis

The genome of strain lai of *L. interrogans* was sequenced by a whole-genome random sequencing method previously applied to other microbial genomes^{3,4,6}. Three different libraries were used in this project. The first two, in pUC18, had inserts of either 1.5–3 kb or 8–10 kb. The third was a 40-kb cosmid library. Altogether, 111,402 sequence reads (Phred value > Q20 (refs 19, 20)) were generated, which gave rise to an overall genome coverage of 8.5 fold, of which 1,600 were from the end sequences of large insert plasmid (8–10-kb) clones and 1,000 were from the end sequences of cosmid clones. The Phred/Phrap/Consed software package^{19–21} was used for quality assessment and sequence assembly. The initial assembly yielded 805 contigs, which were clustered into 145 groups based on linking information from forward and reverse sequence reads. Some contigs were also located on the physical map by Southern analysis. Sequence and/or physical gaps of the chromosomes were closed by primer walking and PCR. The final assembly was checked against the physical map of restriction sites, mapped genes and end sequences of large plasmid and cosmid clones.

Assignment of CDSs

CDSs were determined with Glimmer 2.0 (ref. 22) and the Z-curve method²³, and the results were subjected to further manual inspection. A few CDSs were found by hand curating as guided by BLAST results. BLAST searches against the NCBI non-redundant protein database (or SwissProt, PIR and COG) were performed to determine the similarity. The blast search criteria were as follows: (1) *e*-value = 10^{−5} and (2) at least 60% of the subject sequence was aligned. If there was no database hit, domain analysis was performed by searching the Pfam, PRINTS, PROSITE, ProDom, Block and SMART databases. Transfer RNAs were predicted with tRNAscan-SE²⁴. TopPred²⁵ was used to identify potential membrane-spanning domains in proteins. The presence of signal peptides and the probable position of a cleavage site in secreted proteins were detected with Signal-P. Lipoproteins were identified by scanning for a lipobox ([LV][ASTVI][GAS][C]) in the first 30 amino acids of every protein. Possible metabolic

pathways were examined using the KEGG database¹⁰. Transmembrane helices in proteins were predicted by the THMMH method (Supplementary Information 4). Predicted biological roles were assigned by the classification scheme in ref. 26. In cases in which tertiary structures of hypothetical proteins were predicted, sequences of CDSs were submitted to the SWISS-MODEL server and the illustrations were prepared with Rasmol 2.6.

Deposition of data

In addition to the data deposited at the NCBI database (GB: AE010300 for CI and GB: AE010301 for CII), the *L. interrogans* genome database is also available at <http://www.chgc.sh.cn/lep/> and at <http://bioinfo.hku.hk/LeptoList/>.

BLAST analysis

The BLAST analysis for comparing the best-hit distribution of protein homologues in representative eubacteria with the predicted proteomes of bacteria, virus (phage), archaea and eukarya was based on ref. 27 for studying horizontal gene transfer with modifications. The data were retrieved from NCBI TaxMap (<http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/micr.html>). The CDSs of each bacterium were used in a BLAST search against the database. Only those hits that scored at least 95 bits were collected and ranked. The 'most' similar organism was the one to which the homologous protein bears the strongest similarity with the query CDS.

Enzyme assays

Citramalate synthase activity was assayed as described in ref. 28, with minor modifications. Isopropylmalate synthase activity was assayed as described in ref. 29, with minor modifications.

Received 23 October 2002; accepted 11 March 2003; doi:10.1038/nature01597.

1. Faine, S., Adler, B., Bolin, C. & Perolat, P. *Leptospira and Leptospirosis* (Medisci, Melbourne, 1999).
2. Kmety, E. & Dikken, H. *Classification of the species Leptospira interrogans and history of its serovars*. (Groningen Univ. Press, The Netherlands, 1993).
3. Fraser, C. M. *et al.* Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**, 375–388 (1998).
4. Fraser, C. M. *et al.* Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390**, 580–586 (1997).
5. Saier, M. & Garcia-Lara, J. *The Spirochetes: Molecular and Cellular Biology* (Horizon Scientific Press, Wymondham, 2001).
6. Heidelberg, J. F. *et al.* DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477–483 (2000).
7. Bourhy, P. & Saint Girons, I. Localization of the *Leptospira interrogans metF* gene on the CII secondary chromosome. *FEMS Microbiol. Lett.* **191**, 259–263 (2000).
8. Herter, S. *et al.* Autotrophic CO₂ fixation by *Chloroflexus aurantiacus*: Study of glyoxylate formation and assimilation via the 3-hydroxypropionate cycle. *J. Bacteriol.* **183**, 4305–4316 (2001).
9. Charon, N. W. & Goldstein, S. F. Genetics of motility and chemotaxis of a fascinating group of bacteria: The spirochetes. *Annu. Rev. Genet.* **36**, 47–73 (2002).
10. Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
11. Motaleb, M. A. *et al.* *Borrelia burgdorferi* periplasmic flagella have both skeletal and motility functions. *Proc. Natl Acad. Sci. USA* **97**, 10899–10904 (2000).
12. Picardeau, M., Brenot, A. & Saint Girons, I. First evidence for gene replacement in *Leptospira spp.* Inactivation of *L. biflexa flaB* results in non-motile mutants deficient in endoflagella. *Mol. Microbiol.* **40**, 189–199 (2001).
13. Werts, C. *et al.* Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nature Immunol.* **2**, 346–352 (2001).
14. Arai, H. *et al.* Platelet-activating factor acetylhydrolase (PAF-AH). *J. Biochem. (Tokyo)* **131**, 635–640 (2002).
15. Tuckwell, D. Evolution of von Willebrand factor A (VWA) domains. *Biochem. Soc. Trans.* **27**, 835–840 (1999).
16. Rodrigo, L., Mackness, B., Durrington, P., Hernandez, A. & Mackness, M. Hydrolysis of platelet-activating factor by human serum paraoxonase. *Biochem. J.* **354**, 1–7 (2001).
17. Matsushita, O. *et al.* Gene duplication and multiplicity of collagenases in *Clostridium histolyticum*. *J. Bacteriol.* **181**, 923–933 (1999).
18. Wolf, Y., Rogozin, I., Grishin, N., Tatusov, R. & Koonin, E. V. Genome trees constructed using five different approaches suggest new major bacterial clades. *BMC Evol. Biol.* **1**, 8 (2001).
19. Ewing, B. *et al.* Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185 (1998).
20. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194 (1998).
21. Gordon, D., Abajian, C. & Green, P. Consed: A graphical tool for sequence finishing. *Genome Res.* **8**, 195–202 (1998).
22. Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636–4641 (1999).
23. Zhang, C. T. & Wang, J. Recognition of protein coding genes in the yeast genome at better than 95% accuracy based on the Z curve. *Nucleic Acids Res.* **28**, 2804–2814 (2000).
24. Lowe, T. & Eddy, S. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
25. Claros, M. & von Heijne, G. TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* **10**, 685–686 (1994).
26. Riley, M. Functions of the gene products of *Escherichia coli*. *Microbiol. Rev.* **57**, 862–952 (1993).
27. Olendzenski, L., Liu, L., Zhaxybayeva, O., Murphey, R., Shin, D. G. & Gogarten, J. P. Horizontal transfer of archaeal genes into the *Deinococcaceae*: Detection by molecular and computer-based approaches. *J. Mol. Evol.* **51**, 587–599 (2000).
28. Howell, D. M., Xu, H. & White, R. H. (R)-Citramalate synthase in methanogenic *Archaea*. *J. Bacteriol.* **181**, 331–333 (1999).

29. Kohlhaw, G. B., Leary, T. R. & Umbarger, H. E. Alpha-isopropylmalate synthase from *Salmonella typhimurium*. Purification and properties. *J. Biol. Chem.* **244**, 2218–2225 (1969).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements We thank L. Bao, B.-M. Dai, J. Yan, C. Werts, M. Picardeau and G. Baranton for suggestions and comments on our research strategy and manuscript preparation; C. Jin and G.-C. Liu of the Institute of Microbiology, Chinese Academy of Science, for help in the attempt at assaying the enzymatic activity of PafAH; Y. Liu and H.-G. Zhu for help in preparing the drawings; X. Mao and G. Cai for help in computer simulation; B.-Y. Hu and Y.-X. Nie for help in bacterial culture preparation; and the members of CHGCS for support and encouragement. This work was supported by the National Natural Science Foundation of China, the Chinese National High Technology Development Program (863), the National Key Program for Basic Research (973) and the Sciences and Technology Commission of the People's Government of Shanghai Municipality. It was also supported by the Pôle Sino-Français en Sciences du Vivant et en Génomique and le Programme de Recherches Avancées Franco-Chinois PRA B00-05.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.-P.Z. (e-mail: gpzhao@sibs.ac.cn). The sequences have been submitted to NCBI under accession numbers GB: AE010300 and GB: AE010301 for the large (CI) and small (CII) chromosomes, respectively.

The methylated component of the *Neurospora crassa* genome

Eric U. Selker^{*}, Nikolaos A. Tountas^{†‡}, Sally H. Cross^{†‡}, Brian S. Margolin^{*‡}, Jonathan G. Murphy^{*}, Adrian P. Bird[†] & Michael Freitag^{*}

^{*} Department of Biology and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA

[†] Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, EH9 3JR, UK

Cytosine methylation is common, but not ubiquitous, in eukaryotes. Mammals¹ and the fungus *Neurospora crassa*^{2,3} have about 2–3% of cytosines methylated. In mammals, methylation is almost exclusively in the under-represented CpG dinucleotides, and most CpGs are methylated¹ whereas in *Neurospora*, methylation is not preferentially in CpG dinucleotides and the bulk of the genome is unmethylated⁴. DNA methylation is essential in mammals⁵ but is dispensable in *Neurospora*^{3,6}, making this simple eukaryote a favoured organism in which to study methylation. Recent studies indicate that DNA methylation in *Neurospora* depends on one DNA methyltransferase, DIM-2 (ref. 6), directed by a histone H3 methyltransferase, DIM-5 (ref. 7), but little is known about its cellular and evolutionary functions. As only four methylated sequences have been reported previously in *N. crassa*, we used methyl-binding-domain agarose chromatography⁸ to isolate the methylated component of the genome. DNA sequence analysis shows that the methylated component of the genome consists almost exclusively of relics of transposons that were subject to repeat-induced point mutation—a genome defence system that mutates duplicated sequences⁹.

To isolate the methylated component of the *N. crassa* genome, we cleaved genomic DNA with the 5-methylcytosine-sensitive restriction enzyme *Sau3AI* (recognition sequence GATC) so as to leave intact patches of methylated DNA, and then passed it over a methyl-CpG domain (MBD) column, which fractionates according to the

degree of CpG methylation⁸. Bound DNA was eluted with increasing concentrations of salt, and fractions were analysed by Southern hybridizations, probing for an unmethylated sequence (*am*) and previously identified methylated regions ($\zeta - \eta$, Ψ_{63} and ribosomal DNA; Fig. 1). DNA complementary to the *am* probe eluted principally in pool four but trailed through to pool nine. In contrast, $\zeta - \eta$ sequences peaked later, in pool nine, suggesting that the MBD column successfully fractionated *Neurospora* DNA on the basis of methylation. Considering that the MBD does not bind methylated non-CpG sites¹⁰, these findings suggest co-localization of methylated CpG and non-CpG (*Sau3AI*) sites. By this assay, the

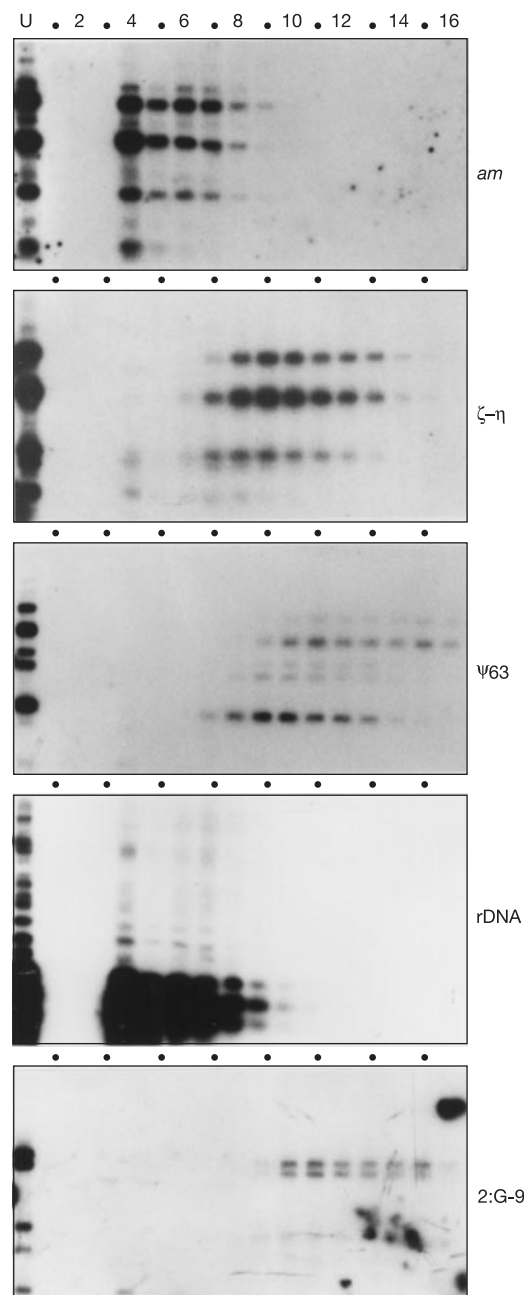


Figure 1 Fractionation of *Neurospora* DNA on a methylated-DNA-binding column. Samples (about 0.5 μ g) of pooled pairs of fractions off the MBD column (1–16 represent column fractions 1–32) were fractionated by agarose gel electrophoresis, along with an unfractionated (U) sample of the *Sau3AI*-digested DNA, blotted to nylon membrane, and probed sequentially for known unmethylated (*am*) and methylated ($\zeta - \eta$, Ψ_{63} and rDNA) sequences, as well as a candidate methylated sequence from this study (2:G-9).

[‡] Present addresses: Center for Cell Signaling, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA (N.A.T.); MRC Human Genetics Unit, Western General Hospital, Edinburgh, EH4 2XU, UK (S.H.C.); Department of Biochemistry and Biophysics, University of California San Francisco, California 94143, USA (B.S.M.).

density of CpG methylation varied within and between sequences. For example, a significant amount of the $\zeta - \eta$ and Ψ_{63} regions of DNA were distributed between pools 4–13, presumably reflecting heterogeneous methylation (Fig. 1). Also, some Ψ_{63} fragments bound more tightly than nearly all the $\zeta - \eta$ DNA, whereas rDNA, which is known to include some methylated cytosines², eluted early, indicating that it has lighter methylation than the two relics of repeat-induced point mutation (RIP).

DNA from tightly bound fractions was cloned and used to probe Southern blots of the MBD-fractionated DNA and/or total genomic DNA digested with *MboI* and *Sau3AI*. Some clones (such as 2:G-9; Fig. 1) represented methylated chromosomal regions as expected, but many showed no evidence of methylation, indicating that the MBD column enriched for, but did not fully purify, the methylated sequences (data not shown). We therefore used differential colony hybridization with probes made from early or late column fractions to screen the library for clones that should represent methylated sequences (see Supplementary Fig. 1). DNA was isolated from about 100 of the resulting candidate methylated DNA clones and tested by probing against DNA from a wild type or from the DNA methyltransferase (*DMTase*) mutant *dim-2* (ref. 6) (Fig. 2). The fraction of high-molecular-mass signal in *Sau3AI* bands was highly variable, consistent with the idea that methylation levels in different chromosomal regions are variable. Results obtained with *dim-2* DNA support the conclusion that the DIM-2 DMTase is responsible for all DNA methylation in vegetative cells of *Neurospora*⁶. About 70% of the clones were found to contain sequences that are unambiguously methylated in the Oak Ridge (OR) wild-type strain (source of the DNA). We thus established a set of experimentally validated methylated DNA clones.

Previous work revealed that trichostatin A (TSA), a potent inhibitor of histone deacetylases, causes hypomethylation of RIP (repeat-induced point mutation) copies of *am*, but does not affect methylation of rDNA and the inactivated transposon *Punt^{RIP1}* in the Ψ_{63} region¹¹. We used our collection of methylated DNA clones to explore the range of the TSA effect and found only minor consequences (Fig. 3). TSA-induced hypomethylation probably reflects the sensitivity of the DIM-5 HMTase to the acetylation state of lysines in histone H3 (H. Tamaru and E.U.S., unpublished data), and the differential effects may reflect different activities of histone acetyltransferases and/or histone deacetylases in various chromosomal regions.

We also used the methylated DNA clones to characterize two mutants defective in methylation (*dim-1* and *dim-3* (refs 12 and 3, respectively)) and to investigate whether the identified methylated sequences are present and methylated in a relatively distant wild-type strain, Mauriceville (M)¹³. Unlike *dim-2*, neither *dim-1* nor *dim-3* caused complete loss of methylation; however, they caused

different patterns of hypomethylation (Fig. 3). Probing of M-DNA revealed many polymorphisms, both in primary sequences and methylation levels, and although many of the sequences are repetitive in both OR and M strains, the levels and patterns of repetition are quite variable. Some of the strongest hybridizing sequences of the OR strain appear absent in the M strain and little or no DNA methylation was observed in the M strain for many sequences (such as 8:A6 and 8:B9). We conclude that regions of DNA methylation are not highly conserved.

The *Neurospora* genome has relatively little (approximately 8%) repetitive DNA¹⁴, presumably because of the operation of RIP⁹. Notably, nearly 50% of the methylated clones hybridized to repetitive DNA (data not shown). We sequenced the ends of 50 methylated clones and calculated two 'RIP indices'¹⁵, based on the observation that RIP preferentially mutates 5'CpA: 5'TpG to generate TpA dinucleotides. For 38 of 50 methylated DNA clones analysed, both indices were diagnostic for RIP. Four clones were positive for one index only and a further three showed evidence of RIP only after detailed analysis (see below). Only 8 of 50 sequences showed ratios typical of unmutated *Neurospora* genes, but three of these were too short to calculate the indices reliably. We conclude that the methylated component of the *Neurospora* genome consists almost exclusively of sequences mutated by RIP.

To gain insight into the natural targets of RIP, we used BLAST¹⁶ to compare the methylated sequences to sequences in GenBank and *N. crassa* genome and proteome databases (including sequences not assembled; <http://www-genome.wi.mit.edu/annotation/fungi/neurospora/>). Notably, 32 of 46 sequences analysed resemble transposons including *Punt^{RIP1}* (Table 1, 5:H8 (ref. 15)), gypsy-like retrotransposon *dab-1* (11:G2 (ref. 17)), the long interspersed nucleotide element (LINE)-like element *Tad* (11:D3 (ref. 18)), and the copia-like element *Tcen* (2:C10, 2:H9, 5:D4, 8:F8 (ref. 19)). We also found relics of transposons not previously known in *Neuro-*

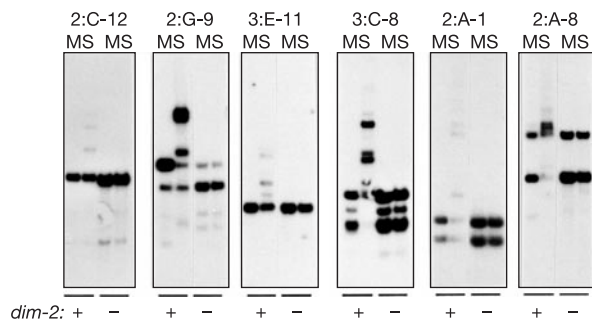


Figure 2 Verification of DNA methylation in genomic DNA regions. Clones representing the indicated potential methylated chromosomal regions were used as probes for Southern blots of *MboI*-digested (M) or *Sau3AI*-digested (S) genomic DNA prepared from wild-type (+) or *dim-2* (–) strains.

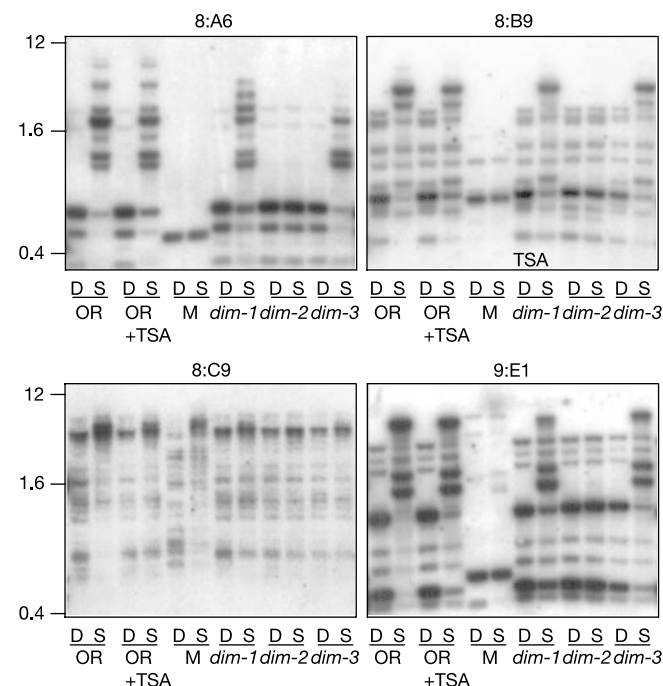


Figure 3 Comparison of DNA methylation in Oak Ridge (OR), Mauriceville (M) and *dim* strains of *N. crassa*. Approximately 1 μ g of DNA from OR (FGSC 2489), M (FGSC 2225), *dim-1* (N2517), *dim-2* (N1104) or *dim-3* (N1089) were digested with *DpnI* (D) or *Sau3AI* (S) and used for Southern hybridizations with probes from clones generated in this study. The OR strain was grown in the presence (+) or absence (–) of 1 μ g ml^{–1} TSA as previously described¹¹. Positions of size standards (kb) are shown on left. Additional examples are presented in Supplementary Fig. 2.

spora, including the *gypsy*-like *Lolligag* (2:A11, 3:C11, Table 1) and DNA-type transposons of the *hAT*²⁰ and *Tc1/mariner*²¹ superfamily (Table 1; see also Supplementary Fig. 3). No relics of true *mariner* transposons occur in the *Neurospora* genome sequence²². Several methylated fragments could not be identified unequivocally as relics of transposons but represented RIP-mutated repeated sequences (for example, 2:A1, 3:E11, 4:D12, 5:F8, 9:E1, 11:E5). We also isolated four methylated segments consisting of RIP-mutated rDNA (7:D8, 7:E5, 7:G9, 8:A10, Table 1) whose origin is uncertain. Finally, three of the four *Tcen* homologues and five additional methylated fragments from centromeric regions provide evidence of DNA methylation in centromeric regions of *Neurospora*.

What caused the methylation of the rare methylated regions with non-diagnostic RIP indices? Previous studies showed that some relics of RIP can induce methylation *de novo* and that methylation can extend far from a sequence serving as a methylation signal²³. Thus, methylation could result from flanking signals. Alternatively, sequences with non-diagnostic RIP indices might nevertheless be products of RIP. Sequence comparisons between the methylated

sequences and related sequences in the genome revealed instances of the latter possibility. Two independently isolated, heavily methylated and (A + T)-rich fragments of a sequence related to the *Minos* transposon in *Drosophila* (8:A6 and 8:B1, Table 1) provide an example. Comparison with a homologous sequence (contig 3.422) revealed 21 and 68 mutations characteristic of RIP, respectively. Similarly, methylated clone 5:A7, which was not predicted to be a product of RIP on the basis of RIP indices (Table 1), contains a RIP-mutated 259-base-pair (bp) segment of the *Dodo2* transposon relic.

To test directly for methylation signals within the 5:A7 region, we introduced 5:A7 segments into *Neurospora* and checked for *de novo* methylation. We targeted single copies of the full 736-bp fragment, the 451-bp *Dodo2* fragment or the adjacent 447-bp sequence, without extraneous (for example, vector) sequences, to the *his-3* or *am* loci and tested for induction of methylation (Supplementary Fig. 4). All fragments induced substantial methylation, indicating that flanking methylation signals are not required. The non-*Dodo2* segment of 5:A7 represents a rare example of *de novo* methylation in

Table 1 *Neurospora crassa* methylated DNA clones

Clone	Accession*	Coordinates	Linkage group†	Methylation (%)‡	Repeats§	TpA ApT	(CpA + TpG) (ApC + GpT)	Comments
2:A1	AY227784	—	—	~50	~2	1.41	0.51	Best match near <i>Cen-III</i>
2:A10	AABX01000180	35973–36621	III (<i>acr-2</i>)	~50	~3	1.28	0.75	Fo <i>Folyt1</i> ; near <i>Cen-III</i> ; Nc <i>Listless</i>
2:A11	AABX01000358	148–2458	VII (<i>wc-1</i>)	~50	~8	1.40	0.36	Mg MGLR-3; Nc <i>Lolligag</i>
2:B3	AABX01000083	8526–9112	I (<i>hsp30</i>)	~60	1	1.26	0.91	Putative RIP-mutated kinase gene (5:D1)
2:C9	AABX01000208	141257–141617	IV (<i>aod-1</i>)	~90	~10	1.14	1.05	Nc <i>Punt</i> linked to Nc <i>Punt</i> ^{RIP1}
2:C10	AABX01000326	36860–37830	VI (<i>cax</i>)	~90	~30	1.47	0.42	Nc <i>Tcen</i> ; near TEL VI L
2:C11	AY227785	—	—	~50	~15	1.27	0.64	Match adjacent to <i>Ant1</i> ; Nc <i>Dodo2</i>
2:C12	AABX01000364	11925–13203	VII (<i>arg-10</i>)	~30	~12	1.38	0.63	Mg Occan; Nc <i>Punt3</i>
2:E8	AABX01000505	20022–21060	—	~50	~12	1.42	0.35	Mg Occan; Nc <i>Punt3</i>
2:G9	AABX01000238	15504–16798	IV (<i>nit-4</i>)	>90	~5	1.33	0.66	Hs <i>Tigger</i> ; Nc <i>Nogo</i>
2:G12	AABX01000227	20529–21108	IV (<i>pyr-1</i>)	~80	~5	1.33	0.74	Hs <i>Tigger</i> ; Nc <i>Nogo</i>
2:H9	AABX01000304	27426–28777	V (<i>ilv-2</i>)	Trace	~30	1.13	0.90	Nc <i>Tcen</i> ; near <i>Cen-V</i>
3:C11	AL670001	10535–12971	II (<i>aro-1</i>)	~70	~8	1.47	0.32	Mg MGLR-3; Nc <i>Lolligag</i> (3:C8)
3:E11	AABX01000368	44939–45532	VII (<i>met-7</i>)	~40	1	1.45	0.89	Adjacent to repeats
4:D12	AY227786	—	—	~30	~10	1.23	0.45	Related to RIP-mutated sequences
5:A7	AABX01000101	12231–12478	I (<i>met-6</i>)	~50	~25	0.75	1.15	An <i>Ant1</i> ; links two contigs; Nc <i>Dodo2</i>
	AABX01000060	1–477	—	—	—	—	—	—
5:B8	AL355932	41108–43653	II (<i>gpd-1</i>)	>90	~5	1.42	0.59	Hs <i>Tigger</i> ; Nc <i>Nogo</i> (8:E12)
5:C6	AABX01000442	2357–3354	—	~5	1	0.76	1.06	Adjacent to Mg <i>pth11</i> homologue
5:C9	AABX01000651	755–1022;1–268	—	~40	~20	1.36	0.61	Adjacent to An <i>Ant1</i> -like transposon
5:D1	AABX01000083	8059–9112	I (<i>hsp30</i>)	~40	1	1.40	0.93	Near <i>Cen-I</i> (2:B3)
5:D2	AABX01000425	2242–3293	—	~50	~30	1.08	0.94	Resembles Nc <i>En/Spm</i> flanks; <i>SNF2</i> -like
5:D4	AABX01000270	3332–4249	V (<i>ilv-2</i>)	~20	~30	1.35	0.54	Nc <i>Tcen</i> ; near <i>Cen-V</i>
5:D6	AABX01000110	3898–4202	I (<i>ad-5</i>)	Trace	~25	1.08	0.95	An <i>Ant1</i> ; Nc <i>Dodo2</i>
5:F8	AABX01000051	4623–6648	I (<i>cyt-1</i>)	~50	~9	1.47	0.50	Repeat region
5:H8	AABX01000241	26017–27007	IV (<i>aod-1</i>)	~40	~10	1.12	0.91	Identical to Nc <i>Punt</i> ^{RIP1}
6:A3	AABX01000207	1412–3547	IV (<i>con-10</i>)	~50	~40	1.26	0.67	Adjacent to Aa <i>Quetzal</i> ; Nc <i>Dodo1</i> (8:F5)
6:B8	AABX01000028	82503–86325	I (<i>vma-11</i>)	~5	~12	1.43	0.56	Mg Occan; Nc <i>Punt3</i>
6:D5	AABX01000436	61687–63770	—	30¶	~30	1.31	0.86	Within repeats near <i>Cen</i>
7:D8	AY227787	—	V L	~40	~150	0.61	0.89	rDNA#
7:E5	AY227788	—	V L	~10	~150	1.00	0.86	rDNA#
7:G9	AY227789	—	V L	~15	~150	0.83	1.05	rDNA#
8:A6	AABX01000266	36323–37403	V (<i>spe-2</i>)	~90	~12	0.86	1.27	Dh <i>Minos</i> ; Nc <i>Dodo1</i> (8:B1)
8:A10	AY227790	—	V L	~20	~150	0.92	1.15	rDNA#
8:B1	AABX01000266	34422–37522	V (<i>spe-2</i>)	>90	~12	0.74	1.14	Dh <i>Minos</i> ; Nc <i>Dodo1</i> (8:A6)
8:B9	AABX01000207	61975–62675	IV (<i>con-10</i>)	~70	~11	0.82	0.96	Mg Pot3; Nc <i>Punt2</i>
8:C9	AABX01000206	136650–141576	IV (<i>pho-5</i>)	~20	~25	1.39	0.52	An <i>Ant1</i> ; Nc <i>Dodo2</i>
8:E12	AL355932	41108–43659	II (<i>gpd-1</i>)	~30	~5	1.42	0.60	Hs <i>Tigger</i> ; Nc <i>Nogo</i> (5:B8)
8:F3	AL670001	20545–22136	II (<i>aro-1</i>)	>90	~8	1.44	0.54	<i>gypsy</i> -type relic
8:F5	AABX01000207	1409–3547	IV (<i>con-10</i>)	~40	~40	1.25	0.67	Adjacent to Aa <i>Quetzal</i> ; Nc <i>Dodo1</i> (6:A3)
8:F8	AABX01000436	38371–40249	—	~50	~30	1.56	0.52	Rice helicase homologue; adjacent Nc <i>Tcen</i>
8:F10	AY251480	450914–454187	III (<i>mip-1</i>)	~50	1	1.47	0.37	—
8:G3	AL669988	70281–71810	II (<i>ro-3</i>)	>90	~9	1.21	0.73	Nh-like transposase; Nc <i>Dodo3</i>
9:E1	AABX01000427	71841–73691	—	~90	~4	1.57	0.53	RIP-mutated repeat
11:D3	AABX01000610	2064–2752	—	30¶	~60	1.39	0.30	Nc <i>Tad</i> ; near <i>Cen</i>
11:E5	AABX01000204	110785–116088	IV (<i>ro-1</i>)	~60	~8	1.28	0.72	RIP-mutated repeat
11:G2	AABX01000297	35403–36604	V (<i>spe-3</i>)	~60	~20	1.78	0.47	Nc <i>dab-1</i>

* GenBank accession numbers for contigs (assembly 3 of WICGR; <http://www-genome.wi.mit.edu/annotation/fungi/neurospora/>). Accession numbers beginning with 'AL' refer to MIPS sequences (<http://www.mips.biochem.mpg.de/proj/neurospora/>). Fragments 2:A1, 2:C11, 4:D12, 7:D8, 7:E5, 7:G9 and 8:A10 had no perfect matches in Assembly 3.

† Linkage group and closest genetic marker (in parentheses).

‡ Degree of DNA methylation estimated from Southern hybridizations (data not shown).

§ The number of transposon relic repeats was estimated from database searches. Matches with a cutoff expected value of 1×10^{-4} were counted. Given the variability of RIP this may be a conservative estimate.

|| Best matches (for example, to transposons) are shown (independent overlapping clones are indicated in parentheses). Aa, *Anopheles albimanus*; An, *Aspergillus niger*; Dh, *Drosophila hydei*; Fo, *Fusarium oxysporum*; Hs, *Homo sapiens*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*; Nh, *Nectria haematococca*. Newly described *Neurospora* transposon relics include *Listless*, *Lolligag*, *Nogo*, *Punt2*, *Punt3* and *Dodo1* to *Dodo3*.

¶ Heterogenous methylation (that is, some bands were unchanged, whereas others were >90% methylated).

RIP-mutated copies of *Neurospora* 25S/28S rDNA, perhaps from the rDNA cluster on linkage group V L (similar or identical sequences in clones 7:E5, 7:D8, b2-1, b2-2, 7:D5, 8:A10).

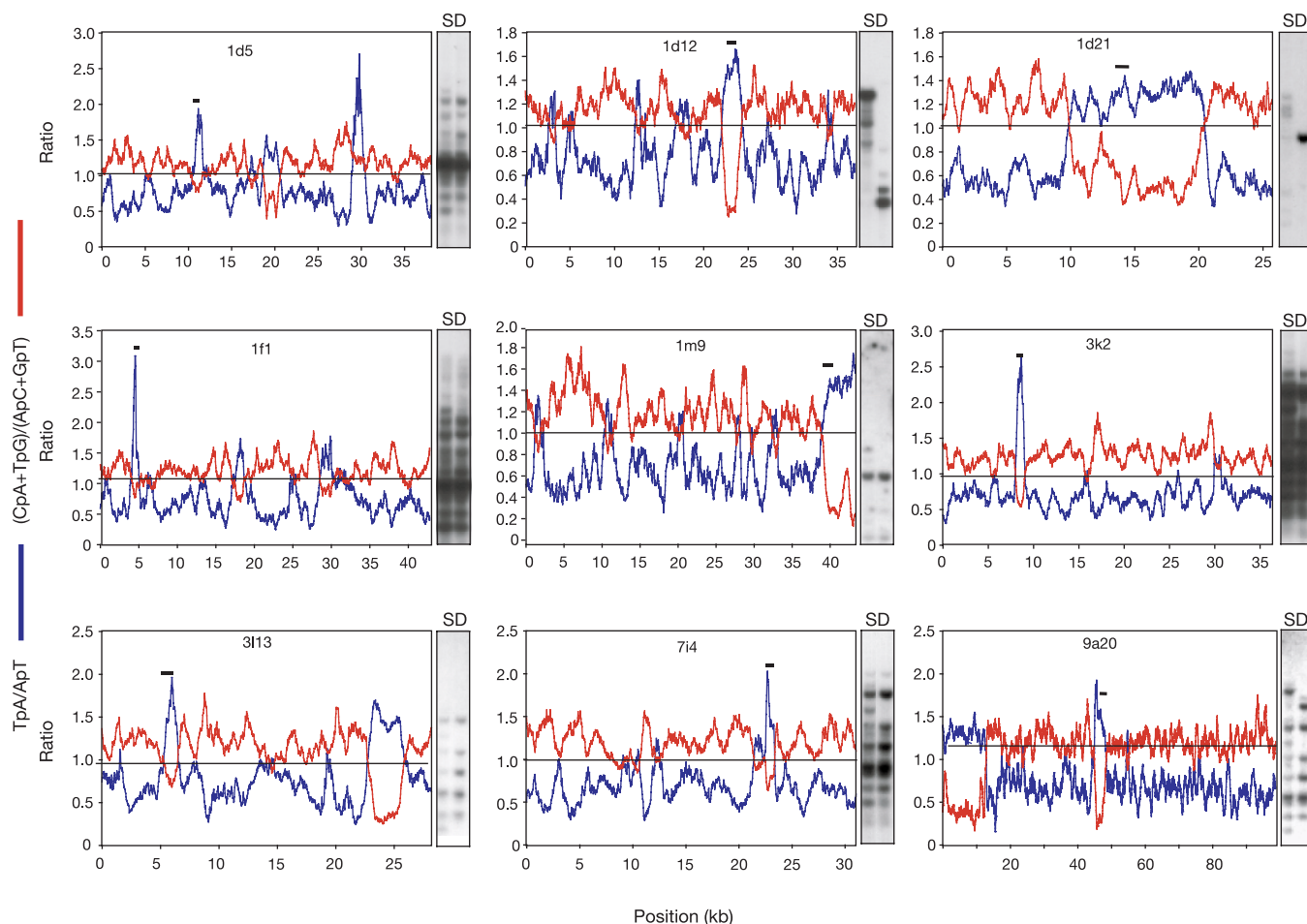


Figure 4 Methylation of sequences with extreme RIP indices. The dinucleotide composition of 3.7 Mb of non-redundant genomic DNA from linkage group V (1d5, 1d12, 1d21, 1f1, 1m9, 7i4 and 9a20) and VI (3k2 and 3l13) (<http://www.mips.biochem.mpg.de/proj/neurospora/>) was analysed in 200 bp windows, with 100 bp shifts in ≈ 40 -kb segments. Sequences with TpA/ApT ratios >0.89 and/or (CpA + TpG)/(ApC + GpT) ratios

<1.03 are considered relics of RIP¹⁵. Representative results from 9/120 contigs analysed are shown. Bars above selected TpA/ApT peaks represent polymerase chain reaction fragments (see Supplementary Table 1) generated to test for methylation of corresponding genomic DNA sequences by Southern hybridization with *Sau3AI*-(S) or *DpnII*-(D) DNA. Resulting autoradiograms are shown to the right of the plots.

Neurospora induced by a sequence showing no evidence of RIP.

Products of RIP are frequently, but not invariably, methylated⁹. To determine whether most natural relics of RIP are methylated, we first calculated RIP indices for sequences of almost an entire *Neurospora* chromosome (Fig. 4). Most (86 of 120) of the approximately 40-kilobase (kb) segments contained at least one apparent relic of RIP. ((TpA/ApT) - (CpA + TpG)/(ApC + GpT) > 0). A total of 174 apparent relics of RIP, ranging from about 0.5 to about 13 kb (average of about 1.5 kb) were found, accounting for approximately 5% of the genomic sequences examined (roughly 0.2 out of 3.7 megabases (Mb)). To determine the fraction of putative relics of RIP that are methylated, we performed Southern hybridizations with probes generated for 20 representative sequences. Notably, 19 of 20 showed evidence of methylation (Fig. 4). Thus DNA methylation and relics of RIP are highly correlated in the *N. crassa* genome (see also ref. 22).

Several model eukaryotes, including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Caenorhabditis elegans* and *Drosophila melanogaster* have no or little DNA methylation. At the other extreme, greater than 25% of cytosines are methylated in DNA of some higher plants²⁴. Mammals and *Neurospora* fall between these extremes. Although the primary function of this modification remains controversial^{25,26}, it is known to prevent gene expression in animals, plants and fungi. Loss of DNA methylation has been shown to reactivate transposons in both *Neurospora*²⁷ and

plants^{28,29}. We show that in *Neurospora*, methylation is found almost exclusively in relics of transposons inactivated by RIP. Thus, all indications are that the primary, and perhaps exclusive, function of DNA methylation in *Neurospora* is to control proliferation of transposons, in conjunction with RIP. Although RIP itself may be limited to certain fungi⁹, methylation is primarily associated with transposons in plants³⁰, raising the possibility that genome defence was the original function of this epigenetic process²⁵. That normal development depends on DNA methylation in some organisms, such as mammals and *Arabidopsis*, may reflect newer roles that methylation has assumed²⁶. □

Methods

Preparation of *Neurospora* DNA

A wild-type strain of *N. crassa* (74-OR23-IVA; Fungal Genetics Stock Center (FGSC) 2489) was grown in stationary liquid Vogel's minimal medium N at 31 °C until saturation (3 days). Genomic DNA was isolated²³ and purified further by phenol/chloroform extractions, CsCl-ethidium bromide equilibrium gradient centrifugation, extraction with isoamyl alcohol and ethanol precipitation. DNA (12 µg) was digested for 2 h with an excess of *Sau3AI*, extracted with phenol:chloroform:isoamyl alcohol (49:49:2), precipitated, and suspended in 40 µl 10 mM Tris (pH 8.0) in preparation for chromatography.

Enrichment of methylated DNA on MBD columns

Neurospora DNA digested with *Sau3AI* (10 µg) was fractionated using an MBD column as described⁸. DNA was ethanol-precipitated from pooled pairs of fractions (2 ml), and aliquots were analysed by gel electrophoresis and Southern hybridization as described²³.

Cloning and identification of methylated sequences

Approximately 25% of fractions 17–18, 19–20, 21–24, 25–28, 29–32 and 33–36 were cloned separately into phosphatase-treated, *Bam* HI-digested pBLUESCRIBE (Stratagene), and approximately 20% of each ligation was electroporated into cells of *Escherichia coli* PKL-F' (*recA*, *lac*, *mcrA*, *mcrB1*, *hsdR2*, *supE44*, *galK2*, *galT22*, *metB1* [*F'* *proAB*, *lacI9Z* Δ M15, Tn10]). Representative white colonies on 5-bromo-4-chloro-3-indoyl- β -D-galactoside (X-Gal) plates were picked (approximately 400 from fractions 21–24; 400 from fractions 25–28; and 100 from each of the other pools) for analysis. Clones were grown in 96-well microtiter dishes with 0.2 ml LB medium supplemented with 10% glycerol and 50 μ g ml⁻¹ ampicillin, and then transferred to -70 °C for storage. Clones representing methylated DNA were tentatively identified by colony hybridization by comparing signals resulting from probing replica blots with labelled DNA from column fractions 7–8 (lane 4 in Fig. 1) and 17–22 (lanes 9–11 in Fig. 1). Purified DNA of promising clones was used to probe *Sau3A*I- or *Mbo* I-digested *Neurospora* DNA by Southern hybridization²³. Inserts were sequenced at the University of Oregon Biotechnology facility.

Computer analyses

To investigate the occurrence of relics of RIP in the genome, we used the Windows program of the GCG Wisconsin Package (Accelrys) to scan 3.7 Mb of linkage group V and VI sequence from the German *Neurospora* genome sequencing consortium (<http://www.mips.biochem.mpg.de/proj/neurospora/>) for hallmarks of RIP, and plotted the data using Excel software (Microsoft). The *Neurospora* genome (<http://www-genome.wi.mit.edu/annotation/fungi/neurospora/>) and GenBank sequences (<http://www.ncbi.nlm.nih.gov/BLAST/>) were searched for matches of 46 sequenced methylated regions (see Table 1) using BLASTN, BLASTX or TBLASTN¹⁶. Alignments were performed using nucleic acid and protein sequences at the Biology Workbench website (<http://workbench.sdsc.edu/>) with ClustalW.

Received 24 December 2002; accepted 14 March 2003; doi:10.1038/nature01564.

1. Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209–213 (1986).
2. Russell, P. J., Rodland, K. D., Cutler, J. E., Rachlin, E. M. & McCloskey, J. A. in *Molecular Genetics of Filamentous Fungi* (ed. Timberlake, W.) 321–332 (Alan Liss, New York, 1985).
3. Foss, H. M., Roberts, C. J., Claeys, K. M. & Selker, E. U. Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation. *Science* **262**, 1737–1741 (1993).
4. Selker, E. U., Fritz, D. Y. & Singer, M. J. Dense non-symmetrical DNA methylation resulting from repeat-induced point mutation (RIP) in *Neurospora*. *Science* **262**, 1724–1728 (1993).
5. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
6. Kouzminova, E. A. & Selker, E. U. *Dim-2* encodes a DNA-methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* **20**, 4309–4323 (2001).
7. Tamaru, H. & Selker, E. U. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414**, 277–283 (2001).
8. Cross, S. H., Charlton, J. A., Nan, X. & Bird, A. P. Purification of CpG islands using a methylated DNA binding column. *Nature Genet.* **6**, 236–244 (1994).
9. Selker, E. U. Repeat-induced gene silencing in fungi. *Adv. Genet.* **46**, 439–450 (2002).
10. Meehan, R. R., Lewis, J. D. & Bird, A. P. Characterization of MeCP2, a vertebrate DNA binding protein that binds methylated DNA. *Nucleic Acids Res.* **20**, 5085–5092 (1992).
11. Selker, E. U. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. *Proc. Natl Acad. Sci. USA* **95**, 9430–9435 (1998).
12. Foss, H. M., Roberts, C. J. & Selker, E. U. Reduced levels and altered patterns of DNA methylation caused by mutations in *Neurospora crassa*. *Mol. Gen. Genet.* **259**, 60–71 (1998).
13. Metzberg, R. L., Stevens, J. N., Selker, E. U. & Morzycka-Wroblewska, E. Identification and chromosomal distribution of 5S rRNA genes in *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* **82**, 2067–2071 (1985).
14. Krumlauf, R. & Marzluf, G. A. Genome organization and characterization of the repetitive and inverted repeat DNA sequences in *Neurospora crassa*. *J. Biol. Chem.* **255**, 1138–1145 (1980).
15. Margolin, B. S. *et al.* A methylated *Neurospora* 5S rRNA pseudogene contains a transposable element inactivated by RIP. *Genetics* **149**, 1787–1797 (1998).
16. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
17. Bibbins, M., Cummings, N. J. & Connerton, I. F. DAB1: a degenerate retrotransposon-like element from *Neurospora crassa*. *Mol. Gen. Genet.* **258**, 431–436 (1998).
18. Cambareri, E. B., Helber, J. & Kinsey, J. A. Tad1-1, an active LINE-like element of *Neurospora crassa*. *Mol. Gen. Genet.* **242**, 658–665 (1994).
19. Cambareri, E. B., Aisner, R. & Carbon, J. Structure of the chromosome VII centromere region in *Neurospora crassa*: degenerate transposons and simple repeats. *Mol. Cell Biol.* **18**, 5465–5477 (1998).
20. Kempken, F. & Windhofer, F. The hAT family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* **110**, 1–9 (2001).
21. Plasterk, R. H., Izsvak, Z. & Ivics, Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet.* **15**, 326–332 (1999).
22. Galagan, J. E. *et al.* The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**, 859–868 (2003).
23. Miao, V. P., Freitag, M. & Selker, E. U. Short TpA-rich segments of the zeta-eta region induce DNA methylation in *Neurospora crassa*. *J. Mol. Biol.* **300**, 249–273 (2000).
24. Paszkowski, J. & Whitham, S. A. Gene silencing and DNA methylation processes. *Curr. Opin. Plant Biol.* **4**, 123–129 (2001).
25. Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
26. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
27. Zhou, Y., Cambareri, E. B. & Kinsey, J. A. DNA methylation inhibits expression and transposition of the *Neurospora* Tad retrotransposon. *Mol. Gen. Genet.* **265**, 748–754 (2001).
28. Miura, A. *et al.* Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* **411**, 212–214 (2001).

29. Singer, T., Yordan, C. & Martienssen, R. A. Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). *Genes Dev.* **15**, 591–602 (2001).
30. Tompa, R. *et al.* Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr. Biol.* **12**, 65–68 (2002).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements We thank F. K. Selker and C. B. Matsen for help with computer analyses; T. Wolfe for technical assistance; G. Mannhaupt for providing some accession numbers; and J. Galagan for comments on the manuscript. E.U.S. acknowledges the hospitality of the Bird laboratory when he initiated this project while on sabbatical, and is thankful to D. Macleod, R. Meehan and F. Antequera for their advice. This work was supported by a US Public Health Service grant to E.U.S. from the National Institutes of Health, and a Senior International Fellowship from the Fogarty International Center of the National Institutes of Health.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.U.S. (e-mail: selker@molbio.uoregon.edu). Sequences are deposited in GenBank. Accession numbers for sequences isolated with the MBD column are listed in Table 1 and accession numbers for the MIPS sequences (Fig. 4) are: AL670542 (1d5), AL451013 (1d12), AL807371 (first 6 kb of 1d21), BX294013 (last 23.55 kb of 1d21), AL513410 (1f1), AL513411 (1m9), BX295538 (3k2), BX295540 (3l13), AL807369 (7i4), BX295539 (9a20).

Cell fusion is the principal source of bone-marrow-derived hepatocytes

Xin Wang*, Holger Willenbring*, Yasmine Akkari*, Yumi Torimaru*, Mark Foster*, Muhsen Al-Dhalimy*, Eric Lagasse†, Milton Finegold‡, Susan Olson* & Markus Grompe*

* Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon 97239, USA

† Stem Cells Inc., Palo Alto, California 94304, USA

‡ Department of Pathology, Texas Children's Hospital, Houston, Texas 77030, USA

Evidence suggests that haematopoietic stem cells might have unexpected developmental plasticity, highlighting therapeutic potential. For example, bone-marrow-derived hepatocytes can repopulate the liver of mice with fumarylacetoacetate hydrolase deficiency and correct their liver disease¹. To determine the underlying mechanism in this murine model, we performed serial transplantation of bone-marrow-derived hepatocytes. Here we show by Southern blot analysis that the repopulating hepatocytes in the liver were heterozygous for alleles unique to the donor marrow, in contrast to the original homozygous donor cells. Furthermore, cytogenetic analysis of hepatocytes transplanted from female donor mice into male recipients demonstrated 80,XXXY (diploid to diploid fusion) and 120,XXXXYY (diploid to tetraploid fusion) karyotypes, indicative of fusion between donor and host cells. We conclude that hepatocytes derived from bone marrow arise from cell fusion and not by differentiation of haematopoietic stem cells.

Recent reports have highlighted the broad developmental potential of bone-marrow-derived stem cells, a phenomenon termed 'stem cell plasticity'. Bone marrow contains haematopoietic stem cells (HSCs)² as well as mesenchymal stem cells³ and multipotent adult progenitor cells⁴. HSCs produce not only all of the blood lineages, but also skeletal muscle⁵, neurons⁶, cardiac muscle⁷, and pulmonary⁸ and liver epithelium^{9–11}. It has been shown that transplantation of HSCs can act as a substitute for hepatocyte transplantation in a murine model of tyrosinaemia, and HSC transplantation can correct this metabolic liver disease¹. Although this indicates that

HSCs might indeed be developmentally plastic and capable of producing both blood cells and hepatocytes, the mechanism underlying this phenomenon remains unclear. Two reports have raised the issue of whether fusion between donor haematopoietic cells and host hepatocytes could explain the apparent stem cell plasticity^{12,13}.

To determine whether the emergence of bone-marrow-derived hepatocytes (BMHs) in fumarylacetoacetate hydrolase (*Fah*)

knockout mice was due to cell fusion, we performed serial transplantation experiments (Fig. 1a, b) using genetically distinct donors and hosts. First, donor alleles were tracked by Southern blot analysis (Fig. 1a). *Fah* mutant mice were lethally irradiated and transplanted with 2×10^6 bone marrow cells from Fanconi anaemia group C gene (*Fancc*) homozygous mutant mice and *Fah* wild-type mice¹⁴. *Fancc* mutant mice have normal livers and were used only to provide a genetic marker traceable by Southern blot. After haematopoietic engraftment, selection for *Fah*-positive (*Fah*⁺) hepatocytes was initiated by withdrawal of the drug 2-(2-nitro-4-trifluoro-methylbenzoyl)-1,3-cyclohexanedione (NTBC), which prevents liver disease in *Fah* mutant mice¹⁵. After 5 months of selection, liver repopulation by BMHs was about 20–30%, as previously reported, and liver function tests were normal (data not shown)¹⁶. Hepatocytes from two independent donors were then serially transplanted into secondary *Fah* mutant recipients. To avoid haematopoietic reconstitution of the secondary host, no preparative irradiation was given. After 3 months of NTBC withdrawal, hepatocytes and haematopoietic tissues were collected. *Fah* immunohistochemistry showed the expected extensive liver repopulation (approximately 50%) but polymerase chain reaction (PCR) analysis failed to detect donor haematopoietic cells in marrow, spleen or thymus (data not shown). Hepatocytes from the secondary recipients were serially transplanted into the final hosts—non-irradiated mice doubly mutant in *Fah* and *Rag1* (ref. 17). After 3 months of *in vivo* selection in these tertiary hosts, pure hepatocyte suspensions were isolated by collagenase perfusion from six mice (three each from the two original donors). Three genotypes were then analysed quantitatively by Southern blots of DNA from these cells (Fig. 1c). Only the final transplant recipient was *Rag1*^{-/-} (Fig. 1a), and therefore the genotype of this locus could be used to measure the degree of liver repopulation, regardless of whether cell fusion had occurred in the repopulating hepatocytes or not. The average donor *Rag1*^{+/+} contribution was 65% (40–91%, Fig. 1c, d). Next, the hepatocyte DNA was genotyped for the *Fancc* and *Fah* mutant and wild-type alleles. If the BMHs had been generated by differentiation of highly plastic HSCs, all of the original donor genotype (that is, *Fancc*^{-/-}/*Fah*^{+/+}) should be preserved in the repopulating cells. However, the fraction of the *Fah*^{+/+} allele was only 30% (8–70%), and that of the *Fancc*^{-/-} allele only 28% (8–65%), less than half of the expected contribution (Fig. 1c). Therefore, the Southern blot results were incompatible with simple differentiation of donor cells to hepatocytes, and indicated cell fusion in all six animals analysed.

To confirm these findings and to ensure that the results were not peculiar to *Fancc* mutant HSCs, bone marrow from female *Fah* wild-type *lacZ* transgenic mice (*Rosa-26*) was transplanted into lethally irradiated male *Fah* mutant recipients, and selection for *Fah*⁺ BMHs was carried out (Fig. 1b). After 5 months, hepatocytes were isolated and analysed in three independent animals. β -Galactosidase staining of single cell preparations revealed hepatocyte replacement levels of at least 20% (Table 1). Hepatocytes were then cultured and metaphase chromosomes were studied. Because *Fah*⁺ cells have a selective survival advantage compared with *Fah*-deficient cells, we expected enrichment for BMHs in culture with tyrosine-containing medium¹⁸. The total number of chromosomes was counted and fluorescent *in situ* hybridization (FISH) with X and Y chromosome paints was performed. Controls included cells from a male wild-type mouse as well as a male *Fah* mutant mouse that had undergone 3 months of cycling NTBC withdrawal. In the controls, almost all nuclei had the expected normal male diploid 40,XY or tetraploid 80,XXYY karyotype (Table 1). In contrast, over 30% of metaphases in all mice with BMHs had the karyotypes predicted to result from a fusion between a diploid donor cell and a diploid host cell (80,XXXY) or a diploid blood cell and a tetraploid host hepatocyte (120,XXXXYY) (Table 1 and Fig. 2). Furthermore, many other aneuploid karyotypes consisting of various combinations of numbers of autosomes and sex chromosomes were

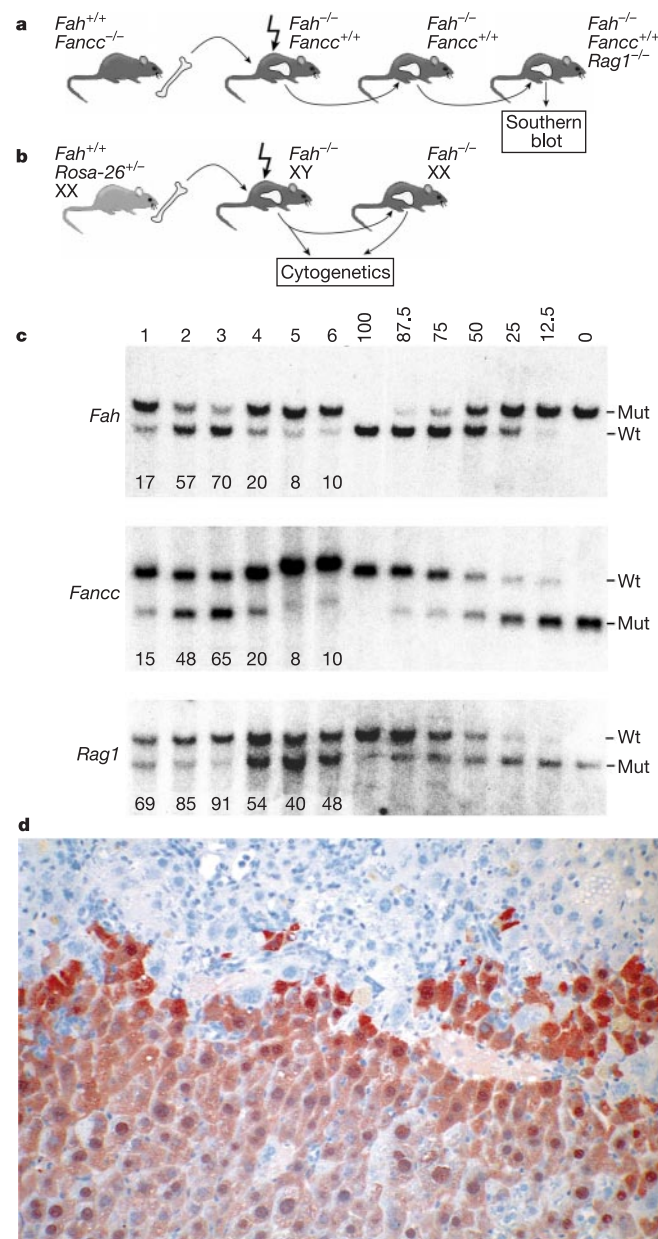


Figure 1 Serial transplantation of bone-marrow-derived hepatocytes. **a**, Bone marrow from a *Fancc* mutant mouse was transplanted into a lethally irradiated *Fah* mutant. BMHs were then serially transplanted into a secondary *Fah* mutant and a tertiary *Fah/Rag1* double mutant mouse. **b**, Bone marrow from a female *Rosa-26* transgenic mouse was transplanted into a lethally irradiated male *Fah* mutant. Hepatocytes were then serially transplanted into a female *Fah* mutant. **c**, Southern blot analysis of hepatocyte DNA from six tertiary serial transplant recipients (*Fah*^{-/-}, *Rag1*^{-/-}). Genomic probes for either *Fah*, *Fancc* or *Rag1* were used and standards with mixes ranging from 100% to 0% of wild-type DNA are shown on the right. The fraction of the original bone marrow donor alleles in the DNA is given in per cent below each lane. In all six samples, the genotype of the original donor bone marrow (*Fah*^{+/+}, *Fancc*^{-/-}) is under-represented. Mut, mutant; Wt, wild type. **d**, *Fah* immunohistochemistry (brown) of a tertiary transplant recipient with approximately 60% liver repopulation by BMHs.

Table 1 Metaphase cytogenetics of cultured hepatocytes

	Male wild-type	Male <i>Fah</i> ^{-/-}	Female BM to male*	Female BM to male*	Female BM to male to female†	Female BM to male to female†
β-gal ⁺ cells (%)	0	0	18	21	32	28
<i>Fah</i> ⁺ cells (%)	100	0	29	41	45	41
Number of metaphases analysed	50	50	50	50	50	65
Normal male (%)						
40,XY	84	52	28	0	8	28
80,XXYY	14	14	18	8	32	12
Cell fusion (%)						
~ 80,XXXYY	0	0	18	24	14	11
~ 120,XXXXYY	0	0	14	36	6	9
X only (40,XX)	0	0	6	6	6	15
Aneuploid (%)‡	2	34	16	26	34	25
Examples§	57,XXY	Hypodiploid, XY (22%) Hypotetraploid, XXXYY (10%) 67,XY	91,3X3Y 117,4X1Y 160,6X2Y 144,4X4Y 166,4X7Y	68,4X2Y 106,3X3Y 92,3X2Y 113,4X1Y 159,6X2Y	48,1X2Y 60,4Y 78,4X2Y 112,4X3Y 167,5X3Y	35,2X1Y 72,2X1Y 77,4X1Y 109,4X1Y 138,6X2Y

β-Galactosidase (β-gal) and *Fah* immunostaining were carried out on parallel cultures. The percentage of nuclei with karyotypes expected for normal male hepatocytes, hepatocytes generated by cell fusion and cells lacking Y chromosomes are shown. For cells negative for the Y chromosome, the fraction of normal 40,XX karyotypes is indicated in brackets below the total percentage.

*The bone marrow (BM)-derived hepatocytes from two male *Fah* mutants transplanted with female bone marrow were analysed (Fig. 1b).

†Hepatocyte metaphases from two female secondary recipients.

‡The percentage of aneuploid metaphases is given.

§Five examples of the abnormal karyotypes are shown.

found (Table 1). Only about half of the cells had the karyotype of normal male liver; that is, 40,XY or 80,XXYY. Furthermore, the fraction of cells containing X but no Y chromosome was very low (6%, Table 1), showing that the female karyotype of the original donor had been mostly lost. Although metaphase spreads provide accurate chromosome analysis, the method is not readily compatible with immunocytochemistry, and it was therefore not possible to prove directly that the cells with fusion karyotypes were indeed *Fah*⁺. Therefore, sequential *Fah* immunocytochemistry and interphase FISH were performed on the same hepatocyte suspensions used for metaphase analysis. Almost all *Fah*⁺ hepatocytes (18 of 19)

studied also contained a Y chromosome, thus confirming that cell fusion had taken place (Fig. 2).

To strengthen further the cytogenetic evidence for fusion, 300,000 BMHs were serially transplanted into a secondary female mouse *Fah*^{-/-} host (see Fig. 1b). *Fah*-deficient hepatocytes are not capable of liver repopulation^{15,18,19} and normal mouse liver contains approximately 5×10^7 hepatocytes. Therefore, less than 1 out of 100 hepatocytes in the secondary recipients would be male unless growth selection of fusion products containing a Y chromosome had occurred. In the liver of the two secondary transplant recipients analysed, about 40% of hepatocytes were *Fah*⁺ (Table 1), confirming

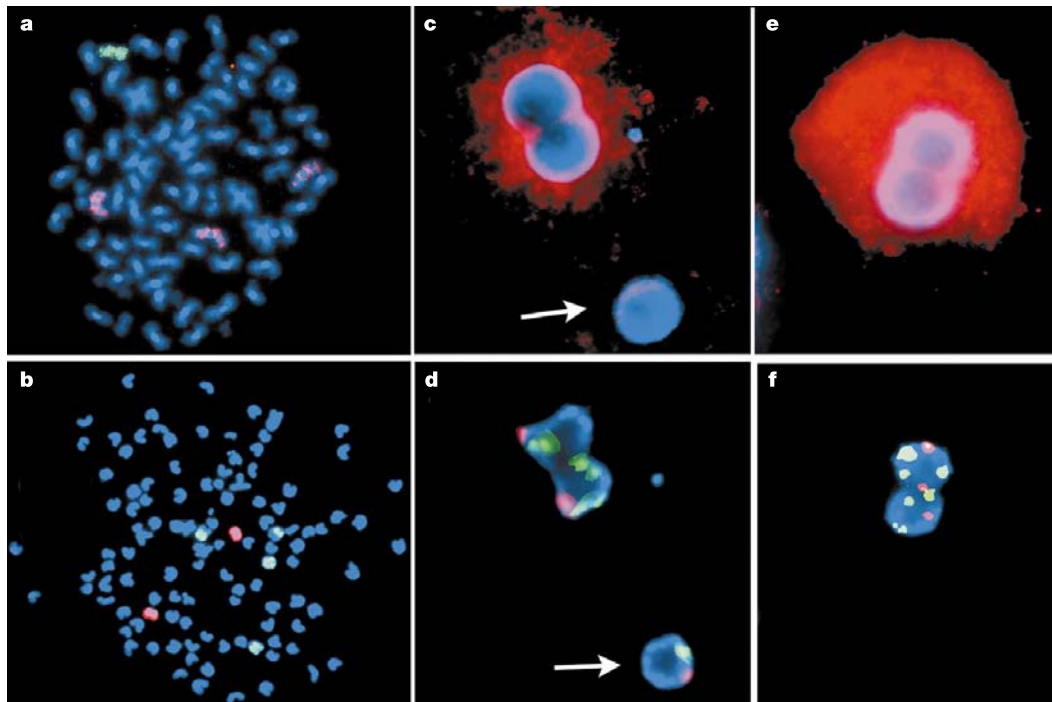


Figure 2 Example of cyto-genetic analysis. **a, b**, Metaphase FISH of cultured bone-marrow-derived hepatocytes from a female-to-male transplant. In **a**, the X chromosomes are pink whereas the Y chromosome is green. In **b**, the Y chromosomes are pink and the X chromosomes are green. Autosomes are stained in blue by DAPI; the Y chromosomes are indicated by arrows. **a**, Fusion karyotype of 80,XXXYY; **b**, fusion karyotype

of 120,XXXXYY. **c-f**, Sequential *Fah* immunocytochemistry and interphase FISH on the same cell. *Fah* staining (red) is shown in **c** and **e**, with the corresponding FISH staining in **d** and **f** below. Y chromosome signals are clearly detectable in *Fah*⁺ hepatocytes. In **c** and **d**, the arrow marks a *Fah*-negative recipient hepatocyte with a normal male (XY) karyotype.

repopulation. In both samples, >80% of hepatocytes at metaphase were positive for the Y chromosome (Table 1), indicating that these cells had been positively selected and therefore must be Fah⁺. Again, many karyotypes indicative of cell fusion were found (Table 1). Interphase FISH of Fah⁺ hepatocytes further confirmed the correlation between Fah positivity and the presence of a Y chromosome in the same cell (Fig. 2).

Together our data indicate that most, if not all, of the BMHs repopulating livers of Fah mutant mice are the product of cell fusion between donor blood cells and host hepatocytes. Rare transdifferentiation events could not be completely ruled out, because 40,XX diploid cells (Table 1) were occasionally observed in male livers repopulated by female bone marrow. However, the significance of these cells is uncertain, because they could simply be contaminating donor haematopoietic cells or they could be the products of a reduction division of fusion hepatocytes (see below). We have shown previously that the bone marrow cells responsible for liver repopulation had the same cell surface phenotype as HSCs¹. In addition, HSCs did not engraft and repopulate the liver directly, but haematopoietic engraftment occurred first, followed by liver repopulation starting about two months later¹⁶. Thus, the kinetics of the emergence of BMHs suggests that the donor cell responsible for fusion is not the haematopoietic stem cell itself, but one of its differentiated progeny, possibly macrophages or T- or B cells. The frequency of hepatocyte replacement was not increased by liver injury¹⁶, indicating that the fusion process occurs stochastically and not as a tissue repair response.

The cytogenetic analysis of BMHs suggests that the fusion probably occurred between a female Fah⁺ diploid donor cell and either diploid or tetraploid male Fah⁺ hepatocytes, resulting in 80,XXXY or 120,XXXXYY hybrids. Haematopoietic cells do not express Fah²⁰ and therefore growth selection of the fusion products cannot take place until the Fah^{+/+} donor nucleus has been reprogrammed to produce Fah. Although the karyotypes of some hybrids were compatible with normal chromosome segregation, many were aneuploid. It is currently unclear whether this is the result of genomic instability inherent to Fah mutant hepatocytes or the result of aberrant chromosome segregation in hybrids. Notably, a high percentage (about 40%) of normal male non-fusion karyotypes (40,XY and 80,XXYY) were found in the female serial transplant recipients (Table 1). This finding suggests that some tetraploid hybrids underwent a reduction division, perhaps by multipolar mitosis²¹, resulting in diploid daughter cells. The 80,XXYY hepatocytes would then be generated by normal polyploidization²². This interpretation is strengthened by the presence of roughly twice as many Fah⁺ as β -galactosidase⁺ BMHs (Table 1). In a hypothetical reduction division, only about half of the Fah⁺ daughter cells would be expected to inherit the lacZ marker because Rosa-26 donor marrow is heterozygous for the lacZ gene²³.

Although our data represent only one example of putative developmental plasticity in haematopoietic stem cells, other examples of this phenomenon could also be due to *in vivo* cell fusion^{5–11}. Therefore, it seems questionable whether HSCs really have any differentiation potential other than the blood lineages ascribed to them traditionally. Although the frequency of spontaneous fusion resulting in BMHs is very low¹⁶, it should nonetheless be noted that these cells can have therapeutic use when present in sufficient quantity¹. At present, this requires a strong selective growth advantage, but it is conceivable that induced cell fusion may achieve the efficiency necessary for the treatment of genetic diseases. □

Methods

Mouse strains and animal husbandry

The following mouse strains inbred on the 129S4 background were used: Fah mutant²⁴, Fancx mutant¹⁴, Rag1 mutant¹⁷ and Escherichia coli lacZ transgenic (Rosa-26)²³. Fah null mice were maintained on NTBC in drinking water²⁵.

Transplantation

Bone marrow collection and transplantation into Fah mutant mice was performed as previously described^{14,16}. Preparative whole-body irradiation consisted of a split dose of 7 Gy ($\times 2$). NTBC was stopped 3 weeks after transplant and then cycled until weight was stabilized¹⁶.

For serial transplantation¹⁹, hepatocytes were isolated by the standard two-step collagenase perfusion²⁶ and enriched to >95% purity by differential centrifugation²⁷. A total of 300,000 viable hepatocytes were injected intrasplenically¹⁹ into Fah^{-/-} or Fah^{-/-}/Rag1^{-/-} recipients as described¹⁵ and NTBC was stopped immediately after transplantation to allow selection of Fah⁺ cells.

Southern blot analysis

For Southern blot analysis²⁸ the following restriction digest/probe combinations were used: Fah: HindIII digest probed with a EcoRI genomic fragment from intron 6 (ref. 24); Fancx: BamHI digest probed with the EcoRI/BamHI genomic fragment containing exon 8 (ref. 14); Rag1: BamHI plus NcoI double-cut-hybridized with probe B-X¹⁷. A Beckman SI Phosphorimager was used to quantify relative band intensities. Signals were compared with genotype standards generated by mixing known ratios of pure wild-type and transgenic homozygous DNA.

Fah and β -galactosidase staining

Hepatocytes were plated on LabTek II CC2 chamber slides (Nalgene) at a density of 10^4 cells cm⁻² in medium consisting of low glucose DMEM (Gibco) supplemented with 10% fetal calf serum (FCS) (Hyclone), 100 ng ml⁻¹ epidermal growth factor, $1 \times$ insulin/transferrin/selenium A (all Gibco). Ten hours after plating, slides were washed with PBS and fixed in methanol at -20 °C for 5 s and in acetone at 4 °C for 3 $\times$ 5 s. Detection of Fah was carried out at room temperature. First, slides were incubated for 20 min with PBS containing 0.05% Tween 20 (Sigma-Aldrich) and 3% normal goat serum (Vector). Then, slides were incubated with a polyclonal rabbit antibody against rat Fah²⁹ diluted 1:10,000. After washing, slides were incubated with alkaline-phosphatase-conjugated swine anti-rabbit antibody (Dako) diluted 1:25. For the detection of antibody-enzyme complexes, the Fast Red Substrate System was used as described by the manufacturer (Dako). For sequential Fah staining plus sex chromosome FISH, the slides were first stained for Fah, and images (bright-field plus fluorescent) of cleanly isolated Fah⁺ cells were captured and their position was recorded. This was followed by the FISH procedure described below. For β -galactosidase visualization, 4×10^5 cells were plated in a 100-mm Primaria dish (Falcon) overnight and then stained³⁰.

Cytogenetics

Single hepatocytes were plated at a density of 7×10^3 cm⁻² on 100-mm Primaria dishes (Falcon) in hepatocyte medium. After 42 h, colcemid (Sigma) was added at a final concentration of 150 ng ml⁻¹. After 4–6 h, cells were digested with trypsin, placed in hypotonic medium consisting of 5% FCS and 75 mM KCl, and fixed to slides. The slides were rinsed with PBS, fixed with 3:1 methanol:acetic acid solution for 80 min, submerged in 70% acetic acid for 2 s, and then allowed to dry overnight. The next day, the slides were washed with $2 \times$ SSC at 37 °C for 30 min, and then treated with 10% pepsin (Sigma) diluted in 40 ml of 0.01 M HCl at 37 °C for 20 min. The slides were then washed with PBS for 5 min, placed in post-fixation solution (1 ml 37% formaldehyde, 0.18 g MgCl₂, 39 ml PBS) for 5 min, and dehydrated with 70%, 80% and 100% ethanol for 2 min each at room temperature.

Immediately before hybridization, the slides were denatured in a 70% formamide solution, pH 7.0 (35 ml formamide, 10 ml $\times 10$ SSC and 5 ml ddH₂O) at 72 °C for 4 min. Meanwhile, paints for the mouse X (FITC) and Y (Cy3) chromosomes (Vysis) were mixed, denatured at 75 °C for 10 min, and preannealed at 37 °C for 30 min. The probes were then added onto the slides and hybridization was carried out using HYBrite (Vysis; 72 °C denaturation temperature, 2 min denaturation time, and reannealing at 37 °C overnight). The next day, the slides were washed with $\times 2$ SSC at 73 °C for 5 min, with $\times 2$ SSC/0.1% NP-40 (Sigma) at room temperature for 1 min, and then counterstained with 125 μ g μ l⁻¹ 4,6-diamidino-2-phenylindole (DAPI) II (Vysis). Cells were observed using a Nikon E800 fluorescence microscope, and captured using CytoVision software from Applied Imaging.

Received 14 November 2002; accepted 25 February 2003; doi:10.1038/nature01531.

Published online 30 March 2003.

1. Lagasse, E. *et al.* Purified hematopoietic stem cells can differentiate into hepatocytes *in vivo*. *Nature Med.* **6**, 1229–1234 (2000).
2. Spangrude, G. J., Heimfeld, S. & Weissman, I. L. Purification and characterization of mouse hematopoietic stem cells. *Science* **241**, 58–62 (1988).
3. Pereira, R. F. *et al.* Cultured adherent cells from marrow can serve as long-lasting precursor cells for bone, cartilage, and lung in irradiated mice. *Science* **276**, 71–74 (1997).
4. Jiang, Y. *et al.* Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. *Exp. Hematol.* **30**, 896–904 (2002).
5. Ferrari, G. *et al.* Muscle regeneration by bone marrow-derived myogenic progenitors. *Hum. Immunol.* **59**, 137–148 (1998).
6. Mezey, E., Chandross, K. J., Harta, G., Maki, R. A. & McKercher, S. R. Turning blood into brain: cells bearing neuronal antigens generated *in vivo* from bone marrow. *Science* **290**, 1959–1962 (2000).
7. Orlic, D. *et al.* Bone marrow cells regenerate infarcted myocardium. *Nature* **410**, 701–705 (2001).
8. Krause, D. S. *et al.* Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* **105**, 369–377 (2001).
9. Petersen, B. E. *et al.* Bone marrow as a potential source of hepatic oval cells. *Science* **284**, 1168–1170 (1999).
10. Theise, N. D. *et al.* Derivation of hepatocytes from bone marrow cells in mice after radiation-induced myeloablation. *Hepatology* **31**, 235–240 (2000).
11. Alison, M. R. *et al.* Hepatocytes from non-hepatic adult stem cells. *Nature* **406**, 257 (2000).

12. Terada, N. *et al.* Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature* **416**, 542–545 (2002).
13. Ying, Q. L., Nichols, J., Evans, E. P. & Smith, A. G. Changing potency by spontaneous fusion. *Nature* **416**, 545–548 (2002).
14. Whitney, M. A. *et al.* Germ cell defects and hematopoietic hypersensitivity to γ -interferon in mice with a targeted disruption of the Fanconi anemia C gene. *Blood* **88**, 49–58 (1996).
15. Overturf, K. *et al.* Hepatocytes corrected by gene therapy are selected *in vivo* in a murine model of hereditary tyrosinaemia type I. *Nature Genet.* **12**, 266–273 (1996).
16. Wang, X. *et al.* Kinetics of liver repopulation after bone marrow transplantation. *Am. J. Pathol.* **161**, 565–574 (2002).
17. Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68**, 869–877 (1992).
18. Overturf, K. *et al.* *Ex vivo* hepatic gene therapy of a mouse model of Hereditary Tyrosinemia Type I. *Hum. Gene Ther.* **9**, 295–304 (1998).
19. Overturf, K., Al-Dhalimy, M., Ou, C. N., Finegold, M. & Grompe, M. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. *Am. J. Pathol.* **151**, 1273–1280 (1997).
20. Mitchell, G. A., Grompe, M., Lambert, M. & Tanguay, R. M. in *The Metabolic and Molecular Basis of Inherited Disease* (eds Scriver, C. R., Beaudet, A. L., Sly, W. & Valle, D.) 1777–1805 (MacGraw-Hill, New York, 2001).
21. Martin, G. M. & Sprague, C. A. Vinblastine induces multipolar mitoses in tetraploid human cells. *Exp. Cell Res.* **63**, 466–467 (1970).
22. Faktor, V. M. & Uryvaeva, I. V. Progressive polyploidy in mouse liver following repeated hepatectomy. *Tsitologia* **17**, 909–916 (1975).
23. Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**, 1513–1523 (1991).
24. Grompe, M. *et al.* Loss of fumarylacetoacetate hydrolase is responsible for the neonatal hepatic dysfunction phenotype of lethal albino mice. *Genes Dev.* **7**, 2298–2307 (1993).
25. Grompe, M. *et al.* Pharmacological correction of neonatal lethal hepatic dysfunction in a murine model of hereditary tyrosinaemia type I. *Nature Genet.* **10**, 453–460 (1995).
26. Berry, M. N. & Friend, D. S. High-yield preparation of isolated rat liver parenchymal cells: a biochemical and fine structural study. *J. Cell Biol.* **43**, 506–520 (1969).
27. Overturf, K., Al-Dhalimy, M., Finegold, M. & Grompe, M. The repopulation potential of hepatocyte populations differing in size and prior mitotic expansion. *Am. J. Pathol.* **155**, 2135–2143 (1999).
28. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (eds Ford, N., Nolan, C. & Ferguson, M.) (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989).
29. Labelle, Y., Puymirat, J. & Tanguay, R. M. Localization of cells in the rat brain expressing fumarylacetoacetate hydrolase, the deficient enzyme in hereditary tyrosinemia type I. *Biochim. Biophys. Acta* **1180**, 250–256 (1993).
30. MacGregor, G. R., Mogg, A. E., Burke, J. F. & Caskey, C. T. Histochemical staining of clonal mammalian cell lines expressing *E. coli* β -galactosidase indicate heterogeneous expression of the bacterial gene. *Somat. Cell Mol. Genet.* **13**, 253–265 (1987).

Acknowledgements This work was supported by NIH grants to M.G., E.L., Y.A. and S.O. X.W. is supported by the Juvenile Diabetes Foundation and H.W. by the Deutsche Forschungsgemeinschaft.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.G. (e-mail: grompem@ohsu.edu).

Transplanted bone marrow regenerates liver by cell fusion

George Vassilopoulos, Pei-Rong Wang & David W. Russell

Division of Hematology, University of Washington, Seattle, Washington 98195, USA

Results from several experimental systems suggest that cells from one tissue type can form other tissue types after transplantation. This could be due to the presence of multipotential or several types of adult stem cells in donor tissues, or alternatively, to fusion of donor and recipient cells. In a model of tyrosinaemia type I, mice with mutations in the fumarylacetoacetate hydrolase gene (*Fah*^{-/-}) regain normal liver function after transplantation of *Fah*^{+/+} bone marrow cells, and form regenerating liver nodules with normal histology that express *Fah*⁺. Here we show that these hepatic nodules contain more mutant than wild-type *Fah* alleles, and that their hepatocytes express both donor and host genes, consistent with polyploid genome formation by

fusion of host and donor cells. Using bone marrow cells marked with integrated foamy virus vectors that express green fluorescent protein, we identify common proviral junctions in hepatic nodules and haematopoietic cells. We also show that the haematopoietic donor genome adopts a more hepatocyte-specific expression profile after cell fusion, as the wild-type *Fah* gene was activated and the pan-haematopoietic CD45 marker was no longer expressed.

Fah-expressing (*Fah*⁺) liver nodules were generated by transplantation of lineage-depleted wild-type bone marrow cells from male mice into lethally irradiated female *Fah*^{-/-} mice, followed by withdrawal of the drug 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (NTBC) to select for growth of *Fah*-expressing hepatocytes¹. Four to five months later, transplant recipients appeared healthy, multiple hepatic nodules were seen on gross examination, and serum bilirubin levels had normalized (Fig. 1a–c). Histological sections revealed regenerating liver nodules consisting of hepatocytes of normal appearance that expressed *Fah*, in a background of dysplastic liver tissue (Fig. 1d–g). Rare *Fah*⁻ hepatocytes were occasionally seen at the periphery of nodules (arrows in Fig. 1g; see also Supplementary Table 1).

We dissected liver nodules from gross specimens where they could be distinguished macroscopically, from tissue blocks where

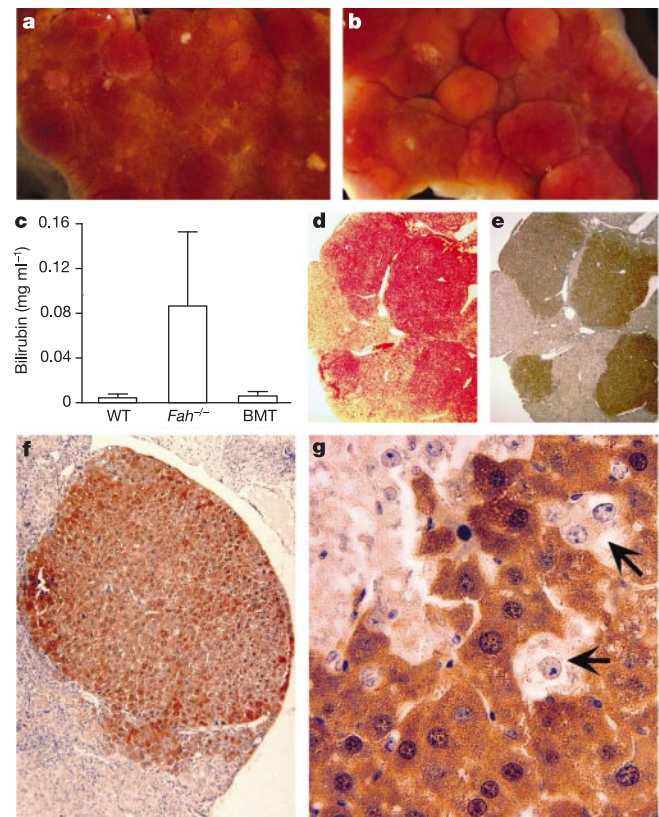


Figure 1 Functional *Fah*⁺ nodules repopulate the livers of *Fah*^{-/-} mice after wild-type bone marrow transplantation. **a**, **b**, Gross liver specimens from transplant recipients showing embedded (**a**) and protruding (**b**) nodules photographed on a dissecting microscope. **c**, Total serum bilirubin levels (mean ± s.d., *n* ≥ 4) from wild-type (WT) mice, *Fah*^{-/-} mice maintained without NTBC for >4 weeks, and *Fah*^{-/-} mice after wild-type bone marrow transplantation (BMT). **d**, **e**, Serial liver sections from transplant recipients stained with haematoxylin and eosin (**d**) or with an anti-*Fah* antibody to stain expressing cells brown (**e–g**). Images were photographed with × 2.5 (**d**, **e**), × 10 (**f**), or × 40 (**g**) objectives. Arrows indicate the locations of non-expressing cells at the edge of an *Fah*⁺ nodule.

adjacent histological sections showed the location of Fah^+ nodules, and from transduced green fluorescent protein-expressing (GFP^+) nodules visualized by fluorescence microscopy (see the secondary transplant described below), taking care to avoid surrounding tissue. Genomic DNA was isolated from each dissected nodule, digested with *HindIII*, and probed for *Fah* sequences to measure the amounts of mutant host (6.9 kilobases) and wild-type donor (5.5 kilobases) alleles² (Fig. 2a). If donor bone marrow cells had transdifferentiated into hepatocytes, or if hepatic stem cells present in the transplanted bone marrow had engrafted in host livers, the regenerating liver nodules should have contained mostly donor DNA. However, 43 of 43 nodules dissected three different ways from eight different mice contained much lower levels of donor DNA than expected (range 12–48%, mean 26%), whereas spleen and bone marrow samples contained the expected high proportion of donor DNA (>90%). Probing for donor Y chromosome sequences confirmed that liver nodules contained low levels of donor DNA ($24.1 \pm 6.6\%$, $n = 11$; Supplementary Fig. 1).

Murine hepatocytes have ploidy values of 2N, 4N, 8N or 16N^{3,4}. Thus fusion of a diploid (2N) haematopoietic cell with a hepatocyte would produce cells with ploidy values of 4N, 6N, 10N or 18N. We calculated the percentage of donor DNA that should be present within regenerating nodules for each of these possibilities by assuming that hepatocytes, Kupffer cells and other remaining cell types accounted for 60%, 15% and 25% of liver cells, respectively^{5,6}, and that phagocytic Kupffer cells were derived from donors^{7–9}. Our studies showed that donor-derived haematopoietic cells normally account for about 10% of liver DNA (Supplementary Fig. 2). Similar calculations were used to predict the percentage of donor

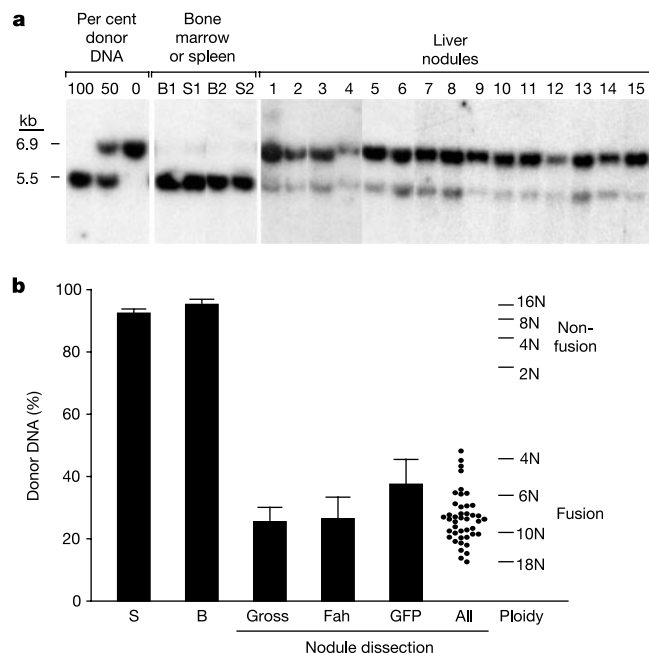


Figure 2 Donor DNA levels in liver nodules. **a**, Southern blot of DNA from bone marrow (B1, B2), spleens (S1, S2) and 15 dissected liver nodules digested with *HindIII* and probed for *Fah* sequences. Controls include wild-type (donor) DNA mixed with *Fah*^{-/-} (host) DNA. **b**, Percentage of donor DNA levels (mean $\pm$ s.d. of wild-type signal/wild-type plus mutant signals) from spleens (S), bone marrow (B) and liver nodules dissected from gross specimens (gross, $n = 16$), *Fah*-stained tissue blocks (*Fah*, $n = 23$), or *GFP*-expressing specimens visualized by fluorescence (*GFP*, $n = 4$). Results from all nodules were also plotted as data points (All, $n = 43$) for a comparison of predicted values from cell fusion and non-fusion models that result in regenerating cells of different ploidy values (see Methods).

DNA values for non-fusion models resulting in donor-derived hepatocytes of the expected ploidy values (2N, 4N, 8N, 16N). The percentage of donor DNA values for every nodule analysed fell within the range expected for fusion events (Fig. 2b), and their distribution suggested that fusion occurred with hepatocytes of all ploidy values, most of which are $\geq 4N$ in mice^{4,10}. As contaminating *Fah*⁻ donor hepatocytes comprised <1% of cells within nodules (Supplementary Table 1), for non-fusion models to produce the observed levels of donor DNA in *Fah*⁺ nodules (mean 26%), about 74% of nodule DNA must have been derived from non-hepatocyte host cells, which is inconsistent with the histology of regenerating liver. As an independent demonstration of cell fusion, we showed through immunofluorescence that the *Fah*⁺ hepatocytes in these nodules express both the donor *Fah* allele and the host allele of $\beta 2$ -microglobulin (Fig. 3).

To determine the source of donor genomes in regenerating liver nodules, we performed secondary transplants of *Fah*^{-/-} mice with wild-type bone marrow cells from a primary transplant recipient

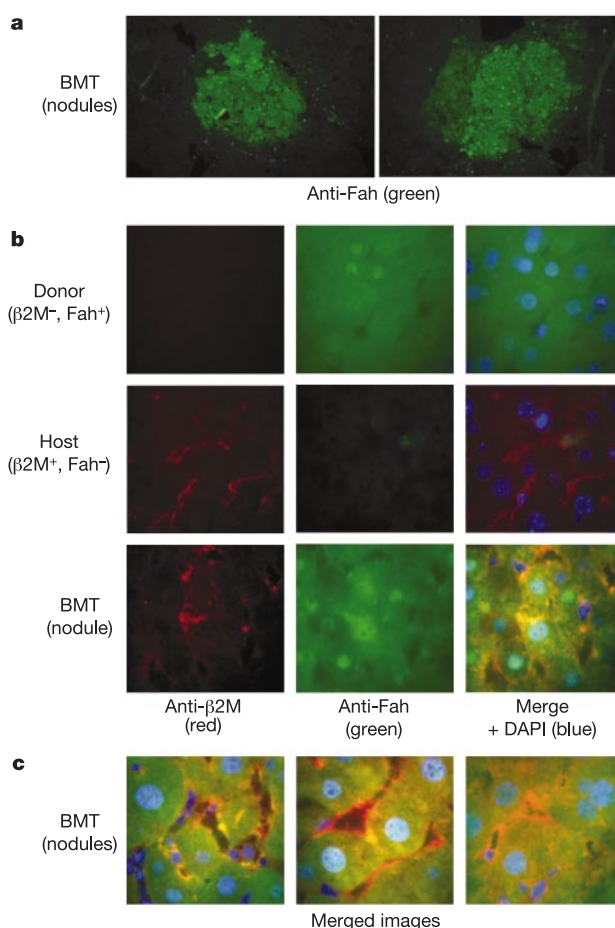


Figure 3 Expression of donor and host genes in nodule hepatocytes. **a**, Low-power fluorescent images of hepatic nodules in bone marrow transplant (BMT) recipients stained green with anti-Fah antibodies. **b**, Individual and merged fluorescent images of liver sections from a wild-type donor (strain 129X1), a *Fah*^{-/-} host (strain C57BL/6) and a regenerating liver nodule in a BMT recipient. Sections were stained with a strain-specific anti- $\beta 2$ -microglobulin ($\beta 2M$) antibody not reactive with the donor strain, anti-Fah antibodies, and 4,6-diamidino-2-phenylindole (DAPI). Host-specific $\beta 2$ -microglobulin can be visualized at the surface of *Fah*⁺ nodule hepatocytes. Each section was photographed with identical exposures. **c**, More merged images prepared as in **b** from other nodules in transplant recipients. Images were photographed with $\times 10$ (**a**) or $\times 100$ (**b**, **c**) objectives.

that had received cells transduced by an integrating foamy retrovirus vector encoding green fluorescent protein (GFP). Under these conditions, secondary recipients partially reconstitute haematopoiesis with monoclonal or oligoclonal donor-derived cells based on provirus junction site analysis¹¹. Two of three transplant recipients died after myeloablation, but one appeared healthy after NTBC withdrawal 4 months after transplantation, and had a normalized serum bilirubin level (0.006 mg ml^{-1}). Peripheral blood from this mouse contained 62% donor (CD45.1⁺) cells, about half of which expressed GFP (Fig. 4a). On histological examination, the nodules appeared normal and were *Fah*⁺ (Fig. 4d, e). GFP⁺ and GFP⁻ liver nodules were present in similar numbers (Fig. 4b, c, f), demonstrating their clonal nature. *Fah*⁺

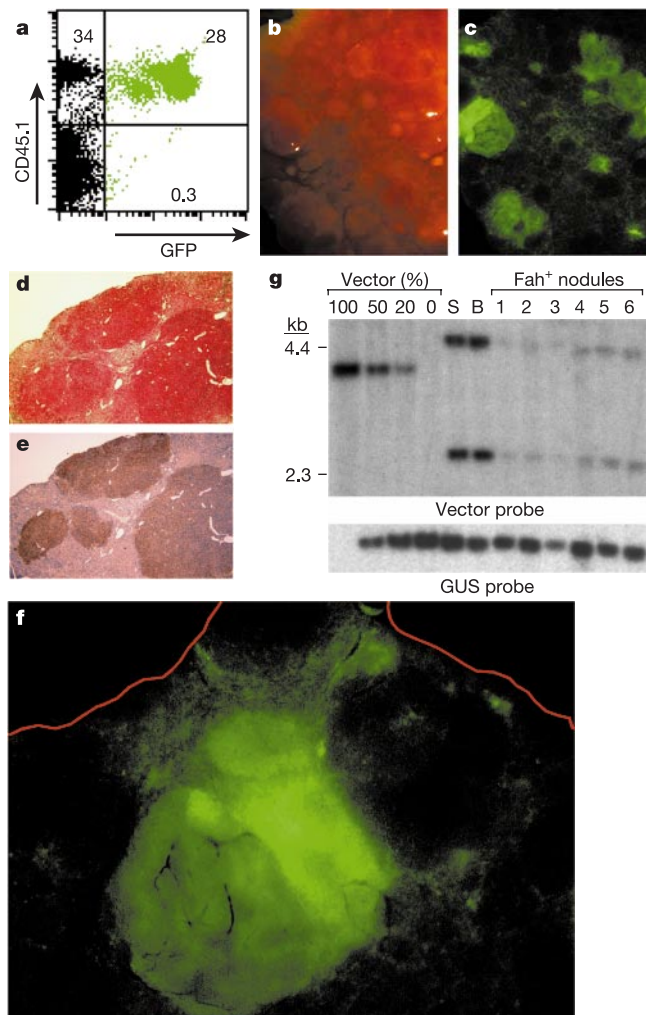


Figure 4 Regenerating liver nodules in a *Fah*^{-/-} mouse after a secondary bone marrow transplant with wild-type cells transduced by a foamy virus GFP vector. **a**, Flow cytometry of the secondary recipient's peripheral blood (donor cells CD45.1⁺). Numbers in quadrants refer to the percentage of cells analysed. **b**, **c**, Identical bright light (**b**) and fluorescent (**c**) dissecting microscope views of the liver. **d**, **e**, Serial liver sections stained with haematoxylin and eosin (**d**) or an anti-*Fah* antibody (**e**) and photographed with a $\times 2.5$ objective. **f**, High-power fluorescent dissecting microscope view of GFP⁺ and GFP⁻ nodules. Red lines indicate liver border. **g**, Southern blot of DNAs from six *Fah*⁺ nodules digested with *Nsi*I and probed for vector sequences. Standards contain the indicated percentage of DNA from human lymphocytes containing a single foamy vector provirus²⁶ mixed with untransduced mouse genomic DNA. S and B indicate spleen and bone marrow samples, respectively. The same blot was reprobed with mouse β -glucuronidase sequences (GUS) to correct for loading.

nodules from tissue blocks and GFP⁺ nodules from the gross specimen were dissected and analysed for *Fah* alleles as in Fig. 2. The percentage values of donor DNA were within the range expected for cell fusion models (range 12–46%, mean 26%, $n = 15$). Provirus junction fragments were identified in DNA samples by cleaving the vector provirus and flanking genomic DNA with *Nsi*I. Bone marrow and spleen samples contained two junction fragments of equal intensity, consistent with haematopoietic reconstitution by the progeny of a single cell containing two randomly integrated vector proviruses (Fig. 4g). The same two fragments were also seen in liver nodule DNA samples, establishing that the donor DNA present in regenerating nodules was derived from haematopoietic cells. The low vector signal in nodules as compared with bone marrow and spleen was consistent with cell fusion models. Some *Fah*⁺ nodules contained more vector DNA than others, reflecting fusion with transduced donor cells (nodules 4–6 in Fig. 4g; average 27% of bone marrow signal), or fusion with non-transduced donor cells, where trace background vector signal presumably was derived from transduced Kupffer cells (nodules 1–3 in Fig. 4g; average 16% of bone marrow signal).

Genetic reprogramming of the haematopoietic donor cell nucleus may be an important step in the formation of normal hepatocytes after cell fusion. To address this issue, we determined which tissues express *Fah* by northern blot analysis (Fig. 5a). Abundant expression was observed in liver and kidney, with lower levels in other tissues. *Fah* messenger RNA was not expressed in haematopoietic spleen and bone marrow tissue (<1% of liver levels). Furthermore, *Fah*⁺ donor-derived haematopoietic cells were not present in the internodule areas of sections stained with anti-*Fah* antibodies (Fig. 5b, and data not shown). Thus, the *Fah* expression detected in regenerating liver nodules represents transcriptional activation of a gene that was silent in the haematopoietic nucleus.

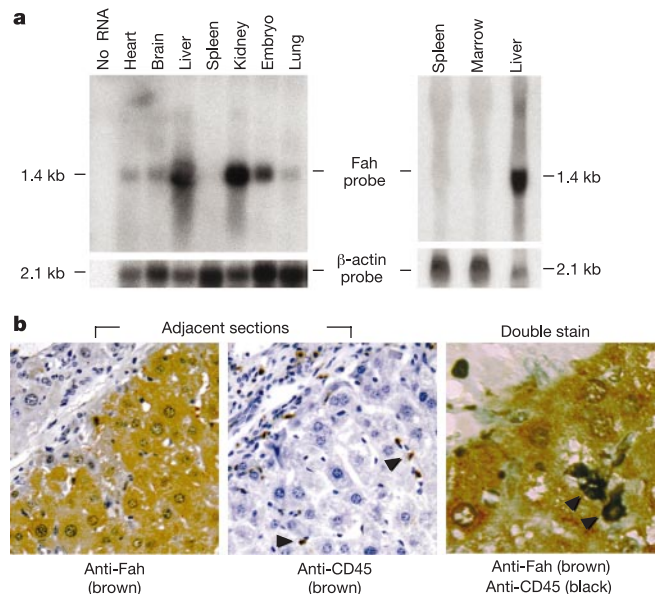


Figure 5 Reprogramming of haematopoietic donor genomes at two loci. **a**, Two northern blots of RNA samples from different wild-type mouse tissues probed with an *Fah* cDNA or β -actin loading control. The left blot was purchased from Ambion, and the right blot contained RNA samples purified from the wild-type donor strain. **b**, *Fah*-expressing and CD45-expressing cells present in the livers of *Fah*^{-/-} mice transplanted with wild-type bone marrow cells were identified by specific antibodies used either in adjacent sections or together in a double stain, as indicated. Arrowheads indicate CD45⁺ haematopoietic cells surrounded by CD45⁻ and *Fah*⁺ hepatocytes. Images were photographed with a $\times 10$ (left and middle panels) or $\times 40$ (right panel) objective.

We also found through histochemical staining that the pan-haematopoietic surface marker CD45 was not expressed in *Fah*⁺ hepatocytes (Fig. 5b), reflecting silencing of a gene that was presumably expressed in the donor cells, and previously shown to be present on the liver-repopulating bone marrow cells in this mouse model¹. Unlike activation of *Fah*, there should not have been selection for silencing of CD45, so this must occur after most fusion events.

Cell fusion occurs during normal mammalian development when mononuclear phagocytes form osteoclasts, or giant cells¹² and myoblasts form myotubes¹³. Transplanted cancer cells can also fuse with host cells *in vivo*^{14,15}, in some cases with cells of haematopoietic origin^{16,17}. In addition, recent studies have shown that spontaneous cell fusion can occur in cultured stem cells^{18,19}. Fusion has not previously been shown to account for *in vivo* findings attributed to possible stem cell plasticity. Our results along with the cytogenetic studies presented in the accompanying paper²⁰ indicate that the regenerating liver nodules in *Fah*^{-/-} transplant recipients are derived from donor haematopoietic cells that fused with host hepatocytes, and not from transdifferentiating haematopoietic stem cells or hepatic stem cells present in bone marrow. Although we cannot determine which haematopoietic cell types fused, phagocytic Kupffer cells are likely candidates given their abundance in the liver, proximity to hepatocytes, and the propensity of macrophages to fuse under other conditions^{12,21,22}. After fusion, the haematopoietic donor genomes underwent reprogramming at the two loci analysed, both activating and silencing genes to resemble hepatocyte expression profiles. These fused cells formed normal-appearing hepatocytes, proliferated, and ultimately corrected the underlying metabolic disorder. □

Methods

Animal care and transplantation

Fah^{-/-} mice² (C57BL/6 strain) were a gift of M. Grompe and were maintained on NTBC as described²³ until transplantation. Bone marrow transplants of irradiated female *Fah*^{-/-} mice were performed with male, lineage-depleted bone marrow cells from *Fah*^{+/+} strain 129X1 (Jackson Laboratory), and transplant recipients were maintained without NTBC, except for 1- or 2-week-long cycles back on NTBC when weight loss occurred, as described¹. Lineage depletion was performed by magnetic separation (Miltenyi Biotec) with antibodies against B220, Mac-1, Gr-1, CD3 and Ter-119 (Pharmingen). For secondary transplantation (Fig. 4) or control experiments to quantify haematopoietic cell contamination levels (Supplementary Fig. 2), primary female C57BL/6 recipients (CD45.2⁺) received lineage-depleted bone marrow from male B6SJL donors (CD45.1⁺) marked with foamy virus vector CGPMscvF (encoding GFP driven from an internal murine stem cell virus promoter), as described¹¹. Seven months later, *Fah*^{-/-} secondary recipients received bone marrow cells from primary recipients. All animal experiments were approved by the Institutional Animal Care Committee.

Tissue analysis

Livers and spleens were removed at the time when mice were killed, and they were used for several purposes. Portions were frozen or paraffin-embedded for histology, then stained with haematoxylin and eosin, with chicken anti-*Fah* antibody (a gift of M. Grompe) and horseradish peroxidase (HRP)-labelled goat anti-chicken antibody, or with rat anti-CD45 antibody (BD Biosciences number 550539) and HRP-labelled rabbit anti-rat antibody. Some sections were counterstained with Mayer's haematoxylin. For immunofluorescence, the anti-*Fah* antibody was detected with fluorescein isothiocyanate-labelled goat anti-chicken antibody, and the mouse anti-β2 microglobulin antibody (BD Biosciences number 555299) was detected with Texas red-labelled horse anti-mouse antibody. Liver nodules were dissected from gross specimens based on their appearance, or in some cases based on GFP expression. *Fah*⁺ nodules were dissected from 2–3-mm-thick tissue blocks after staining an adjacent section for *Fah* expression, by excising a shallow, conical sample from the centre of each nodule. Southern and northern blots were performed by standard techniques and quantified by PhosphorImager, which was accurate to within $3.4 \pm 3.6\%$ on the basis of repeat analysis of standards (not shown). The *Fah* genomic probe² was an *EcoRI*/*HindIII* fragment obtained by polymerase chain reaction of mouse genomic DNA with primers AGTGACCTAGCCATGAACG and GGGGTAGAGCTGGGACTCA. The Y chromosome²⁴, β-glucuronidase (GUS)¹¹, foamy virus vector Gag¹¹, and *Fah* complementary DNA probes²⁵ have been described.

Ploidy calculations

The following equations were used to calculate per cent donor DNA values (Fig. 2b) for fusion models where regenerating hepatocytes were derived from one diploid wild-type cell and a hepatocyte of ploidy 2N, 4N, 8N or 16N, and for non-fusion models where regenerating hepatocytes were all derived from bone marrow and formed nodules of ploidy 2N, 4N, 8N or 16N. For fusion models: per cent donor genomes =

$100((P_B/P_H)H + K)/(H + K + R)$; for non-fusion models: per cent donor genomes = $100(H + K)/(H + K + R)$. The fraction of DNA from hepatocytes (H) is $0.6P_H/(0.6P_H + 0.15P_K + 0.25P_R)$; the fraction of DNA from Kupffer cells (K) is $0.15P_K/(0.6P_H + 0.15P_K + 0.25P_R)$; the fraction of DNA from remaining (R) cells is $0.25P_R/(0.6P_H + 0.15P_K + 0.25P_R)$. P_B , P_H , P_K and P_R indicate the ploidy of bone-marrow-derived cells, hepatocytes, Kupffer cells and remaining cells, respectively. P_B , P_K and P_R are 2N, whereas P_H varied. P_B/P_H is a correction factor for the fraction of donor genomes present in hepatocytes of different ploidy values.

Received 13 November 2002; accepted 12 March 2003; doi:10.1038/nature01539.

Published online 30 March 2003.

- Lagasse, E. *et al.* Purified hematopoietic stem cells can differentiate into hepatocytes *in vivo*. *Nature Med.* **6**, 1229–1234 (2000).
- Grompe, M. *et al.* Loss of fumarylacetoacetate hydrolase is responsible for the neonatal hepatic dysfunction phenotype of lethal albino mice. *Genes Dev.* **7**, 2298–2307 (1993).
- Epstein, C. J. & Gatens, E. A. Nuclear ploidy in mammalian parenchymal liver cells. *Nature* **214**, 1050–1051 (1967).
- Bohm, N. & Noltemeyer, N. Development of binuclearity and DNA-polyploidization in the growing mouse liver. *Histochemistry* **72**, 55–61 (1981).
- Arias, I. M. (eds) *The Liver* (Raven, New York, 1994).
- Zakim, D. & Boyer, T. D. (eds) *Hepatology* (W. B. Saunders Company, Philadelphia, Pennsylvania, 1990).
- Gale, R. P., Sparkes, R. S. & Golde, D. W. Bone marrow origin of hepatic macrophages (Kupffer cells) in humans. *Science* **201**, 937–938 (1978).
- Crofton, R. W., Diesselhoff-den Dulk, M. M. & van Furth, R. The origin, kinetics, and characteristics of the Kupffer cells in the normal steady state. *J. Exp. Med.* **148**, 1–17 (1978).
- Paradis, K., Sharp, H. L., Vallera, D. A. & Blazar, B. R. Kupffer cell engraftment across the major histocompatibility barrier in mice: bone marrow origin, class II antigen expression, and antigen-presenting capacity. *J. Pediatr. Gastroenterol. Nutr.* **11**, 525–533 (1990).
- Weglartz, T. C., Degen, J. L. & Sandgren, E. P. Hepatocyte transplantation into diseased mouse liver. Kinetics of parenchymal repopulation and identification of the proliferative capacity of tetraploid and octaploid hepatocytes. *Am. J. Pathol.* **157**, 1963–1974 (2000).
- Vassilopoulos, G., Trobridge, G., Josephson, N. C. & Russell, D. W. Gene transfer into murine hematopoietic stem cells with helper-free foamy virus vectors. *Blood* **98**, 604–609 (2001).
- Vignery, A. Osteoclasts and giant cells: macrophage-macrophage fusion mechanism. *Int. J. Exp. Pathol.* **81**, 291–304 (2000).
- Mintz, B. & Baker, W. W. Normal mammalian muscle differentiation and gene control of isocitrate dehydrogenase synthesis. *Proc. Natl Acad. Sci. USA* **58**, 592–598 (1967).
- Goldenberg, D. M., Bhan, R. D. & Pavia, R. A. *In vivo* human-hamster somatic cell fusion indicated by glucose 6-phosphate dehydrogenase and lactate dehydrogenase profiles. *Cancer Res.* **31**, 1148–1152 (1971).
- Wiener, F., Fenyo, E. M., Klein, G. & Harris, H. Fusion of tumour cells with host cells. *Nat. New Biol.* **238**, 155–159 (1972).
- Kerbel, R. S., Lagarde, A. E., Dennis, J. W. & Donaghy, T. P. Spontaneous fusion *in vivo* between normal host and tumour cells: possible contribution to tumour progression and metastasis studied with a lectin-resistant mutant tumour. *Mol. Cell Biol.* **3**, 523–538 (1983).
- Wiener, F., Fenyo, E. M. & Klein, G. Tumor-host cell hybrids in radiochimera. *Proc. Natl Acad. Sci. USA* **71**, 148–152 (1974).
- Terada, N. *et al.* Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature* **416**, 542–545 (2002).
- Ying, Q. L., Nichols, J., Evans, E. P. & Smith, A. G. Changing potency by spontaneous fusion. *Nature* **416**, 545–548 (2002).
- Wang, X. *et al.* Cell fusion is the principal source of bone-marrow-derived hepatocytes. *Nature* **422**, 897–901 (2003).
- Parwaresch, M. R., Kreipe, H. & Radzun, H. J. Human macrophage hybrid forming spontaneous giant cells. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **51**, 89–96 (1986).
- Falzone, S. *et al.* The purinergic P2Z receptor of human macrophage cells. Characterization and possible physiological role. *J. Clin. Invest.* **95**, 1207–1216 (1995).
- Grompe, M. *et al.* Pharmacological correction of neonatal lethal hepatic dysfunction in a murine model of hereditary tyrosinaemia type I. *Nature Genet.* **10**, 453–460 (1995).
- Bishop, C. E., Boursot, P., Baron, B., Bonhomme, F. & Hatat, D. Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. *Nature* **315**, 70–72 (1985).
- Grompe, M. & al-Dhalimi, M. Nucleotide sequence of a cDNA encoding murine fumarylacetoacetate hydrolase. *Biochem. Med. Metab. Biol.* **48**, 26–31 (1992).
- Josephson, N. C. *et al.* Transduction of human NOD/SCID-repopulating cells with both lymphoid and myeloid potential by foamy virus vectors. *Proc. Natl Acad. Sci. USA* **99**, 8295–8300 (2002).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements The authors thank M. Grompe for *Fah*^{-/-} mice, anti-*Fah* antibodies and advice; T. Papayannopoulou and E. Skarpidi for review of the manuscript; and C. Xu, Y. Jiang and R. Newton for technical assistance. This work was supported by grants from the US National Institutes of Health.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.R. (e-mail: drussell@u.washington.edu).

Chibby, a nuclear β -catenin-associated antagonist of the Wnt/Wingless pathway

Ken-Ichi Takemaru^{*,†}, Shinji Yamaguchi^{†,‡}, Young Sik Lee[‡], Yang Zhang^{*}, Richard W. Carthew[‡] & Randall T. Moon^{*}

^{*} Howard Hughes Medical Institute, Room K536C Health Sciences Building, Campus Box 357750, Department of Pharmacology, and Center for Developmental Biology, University of Washington School of Medicine, Seattle, Washington 98195, USA

[‡] Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, Illinois 60208-3500, USA

[†] These authors contributed equally to this work

Inappropriate activation of downstream target genes by the oncoprotein β -catenin is implicated in development of numerous human cancers^{1,2}. β -catenin and its fruitfly counterpart Armadillo act as a coactivator in the canonical Wnt/Wingless pathway by binding to Tcf/Lef transcription factors^{3–6}. Here we report a conserved nuclear protein, named Chibby, which was identified in a screen for proteins that directly interact with the C-terminal region of β -catenin. In mammalian cultured cells we demonstrate that Chibby inhibits β -catenin-mediated transcriptional activation by competing with Lef-1 to bind to β -catenin. Inhibition of *Drosophila* Chibby by RNA interference results in segment polarity defects that mimic a *wingless* gain-of-function phenotype, and overexpression of the *wingless* target genes *engrailed* and *Ultrabithorax*. In addition, epistasis experiments indicate that *chibby* acts downstream of *wingless* and upstream of *armadillo*.

β -Catenin/Armadillo (Arm) is composed of 12 Arm repeats flanked by unique N and C termini⁷. Both the N- and C-terminal regions exhibit transcriptional activation activity, but the most potent transactivation domain resides in the C terminus^{8,9}. To isolate cofactors that bind to the C-terminal region of β -catenin and that could modulate its transducing activity, we carried out a screen with the yeast Ras recruitment system (RRS)¹⁰ that detects protein–protein interactions at the inner surface of the plasma membrane. As bait in the screen we used Arm repeat 8 to the C-terminus¹¹. Of 79 positive clones, 59 encoded a full-length previously unknown human protein we named Chibby (Cby; ‘small’ in Japanese, named after the RNAi phenotype in fly; see Fig. 1d). Human Cby contains 126 amino acids with predicted nuclear localization signals and a coiled-coil domain, and is evolutionarily conserved from fly to human (Supplementary Fig. 1).

With regard to the expression of *cby*, northern blot analysis detects a single 1.2-kilobase (kb) RNA that is expressed at different levels in several tissues of the adult human (Fig. 2a). Cby protein is predominantly nuclear in a punctate manner, as visualized by immunostaining of the endogenous protein in COS7 cells and other cell lines (Fig. 2b; data not shown). Transfection of Wnt-1, Wnt-5a or β -catenin did not significantly change the nuclear localization of Cby in COS7 cells (data not shown).

We then used the yeast RRS to define the Cby binding domain of β -catenin and found that Arm repeat 10 to the C-terminus is required for its interaction with Cby (Fig. 3a). Consistent with this finding, *in vitro* pull-down assays demonstrated that Cby directly binds to the C-terminal region of β -catenin (Fig. 3b, lane 6) as well as to the full-length protein (Fig. 3b, lane 9), but not to the N-terminal region (Fig. 3b, lane 3). On the other hand, Cby associates with β -catenin through its C-terminal half (Fig. 3c). To test for their interaction *in vivo*, we co-transfected human embryonic kidney 293T cells with expression plasmids encoding Myc-tagged β -catenin and Flag-tagged Cby, and immunoprecipitated

the lysates. As predicted, β -catenin co-immunoprecipitates with Cby and vice versa (Fig. 3d). Finally, we investigated whether the endogenous proteins interact by immunoprecipitations from nuclear extracts of non-transfected 293T cells. As shown in Fig. 3e, immunoprecipitation of β -catenin pulls down Cby, but not the unrelated nuclear proteins Cyclin A or TFIIB (data not shown). We conclude that Cby is normally present in a complex with β -catenin *in vivo*.

To evaluate whether Cby modifies β -catenin-dependent signaling, we performed Tcf (T-cell factor) reporter assays¹² in 293T cells. Overexpression of Cby repressed the transcriptional activation of the β -catenin-dependent Tcf reporter TOPFLASH by Wnt-1 or stabilized β -catenin¹³ in a dose-dependent fashion, showing minimal effects on the mutant reporter FOPFLASH (Fig. 4a). Similar results were obtained using other cell types including COS7 and NIH3T3 (data not shown). Consistent with these results, the

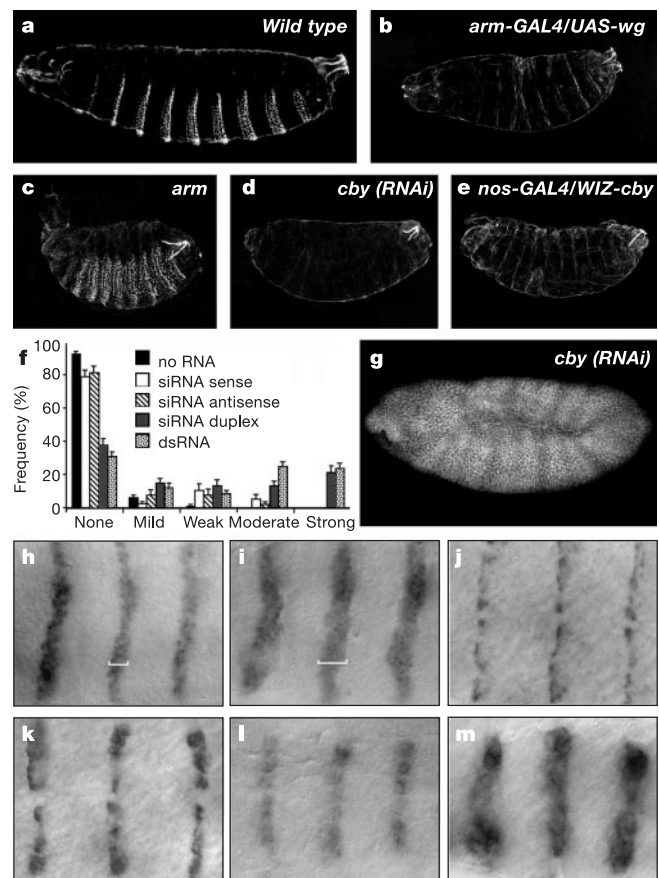


Figure 1 RNAi of *cby* in *Drosophila* embryos causes ectopic activation of the *wg* pathway. Anterior is left and ventral bottom. **a**, A wild-type embryo cuticle. **b**, UAS-*wg*^{ts} expressed ubiquitously at 17 °C under *arm-GAL4* control eliminates denticles. **c**, Loss of *arm* (*arm*^{XP33}) eliminates naked cuticle. *cby* RNAi generated by dsRNA injection (**d**) or by WIZ-*cby* ubiquitously expressed under *nos-GAL4* control (**e**) produce embryos that resemble those ubiquitously expressing *wg*. Additionally, embryos are shorter, lacking head structures. **f**, *cby(RNAi)* phenotypes are similar whether induced by dsRNA or siRNA injections. Phenotypes were categorized as: None, wild-type; Mild, one segment affected; Weak, two to four segments affected; Moderate, five or more segments affected; Strong, completely naked cuticle. The number of treated embryos that fell into each category is shown as frequency. **g**, Abundance and localization of *Arm* protein is not significantly affected in stage 9 *cby(RNAi)* embryos ($n = 40$). **h**, En antibody staining in ps3-5 of a stage 11 embryo. **i**, Injection with *cby* dsRNA expands each stripe (brackets) of En expression. **j**, *wg* RNA distribution in ps3-5 of a stage 11 embryo. **k**, *cby(RNAi)* increases *wg* transcript levels. **l**, Antibody staining for Wg in ps11-13 of a stage 11 embryo. **m**, *cby(RNAi)* expands and enhances Wg protein expression.

activity of the β -catenin-dependent *cyclin D1* reporter¹⁴ was also repressed by co-expression of Cby with stabilized β -catenin (data not shown). We note that β -catenin protein remains stable in the presence of high levels of Cby, indicating that Cby does not promote β -catenin degradation (data not shown).

To investigate possible alternative mechanisms by which Cby inhibits β -catenin-dependent gene activation, we co-transfected Myc-tagged β -catenin and Flag-tagged Cby with increasing amounts of haemagglutinin (HA)-tagged Lef-1 (lymphoid-enhancer factor-1), then immunoprecipitated with an anti-Myc antibody. Because the amount of Cby co-precipitating with β -catenin decreases as the amount of Lef-1 associating with β -catenin increases, we conclude that Cby competes with Lef-1 for binding to β -catenin (Fig. 4b). In contrast, Cby had no apparent effect on the interaction between β -catenin and APC or CBP (data not shown).

To further explore the role of Cby as a possible β -catenin antagonist, we depleted endogenous Cby protein using an antisense morpholino oligo against human Cby (CbyMO) in 293T cells. As Cby protein levels decreased (Fig. 4c), human CbyMO promoted an increase in TOPFLASH activity (Fig. 4d). As a control, we showed that co-transfection with a mouse Cby expression construct lacking the human CbyMO target sequence rescued the TOPFLASH activity to near the basal level (Fig. 4d). Collectively, these data suggest that Cby negatively regulates β -catenin-mediated transactivation by competing with Tcf/Lef transcription factors in mammalian cells.

To investigate the role of Cby in Wnt signalling during development, we turned to *Drosophila* because the fruitfly expresses a Cby orthologue that, like the human Cby described above, inhibits β -catenin-dependent activation of TOPFLASH in mammalian cultured cells (data not shown). We injected *cby* double-stranded RNA into wild-type embryos, which resulted in transformation of ventral denticles into naked cuticle (Fig. 1d compared to control in Fig. 1a). This phenotype is similar to embryos that overexpress *wingless* (*wg*) in all epidermal cells¹⁵ (Fig. 1b). In contrast, losing *wg* pathway activity results in transformation of naked cuticle into denticles¹⁶ (Fig. 1c). In addition, the head structures of *cby*(RNAi) embryos were missing, and the embryos were small. These phenotypes were also observed in embryos that express intron-spliced snapback RNA¹⁷ corresponding to *cby* (Fig. 1e). To confirm the

specificity of the RNAi phenotype, we injected a short interfering RNA¹⁸ whose sequence did not overlap with either the dsRNA or snapback RNA used earlier. The siRNA duplex produced a highly similar phenotype, indicating that the phenotype is caused by specific reduction of *cby* activity (Fig. 1f).

The cuticle phenotype suggests that *cby* functions as an antagonist of the *wg* pathway. To test this possibility further, we examined expression of the gene *engrailed* (*en*), which is transcriptionally activated by *wg* signalling in the cuticle-secreting embryonic epidermis¹⁹. *En* is normally expressed in segmental stripes two cells wide, immediately behind cells expressing *wg* (Fig. 1h). In *cby*(RNAi) embryos, the *En* stripes expanded by another row of cells (Fig. 1i), a phenotype similar to one observed when embryos overexpress *wg* in the *en* domain²⁰. This indicates that *cby* normally

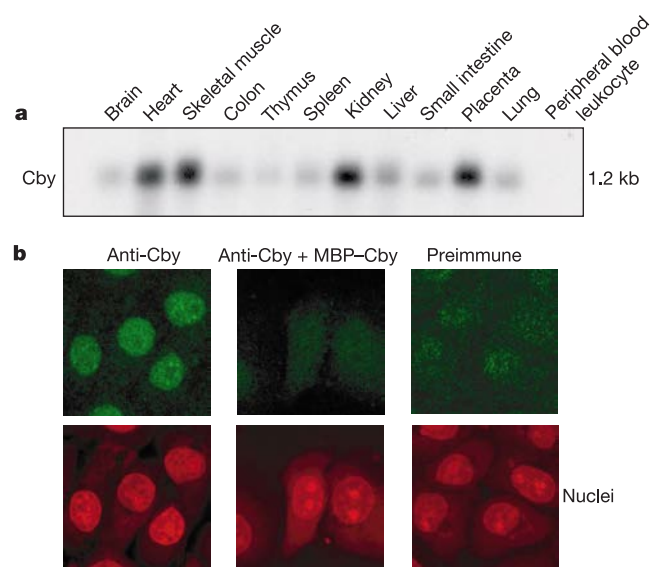


Figure 2 Cby encodes a nuclear protein. **a**, Northern blot analysis of human tissues. **b**, Cby is localized in the nucleus, as stained with an anti-Cby antibody in COS7 cells (green, top row). Preincubation of the Cby antibody with an MBP-Cby fusion protein or the preimmune serum shows weak background staining. Nuclei were stained with BO-PRO-3 (red, bottom row).

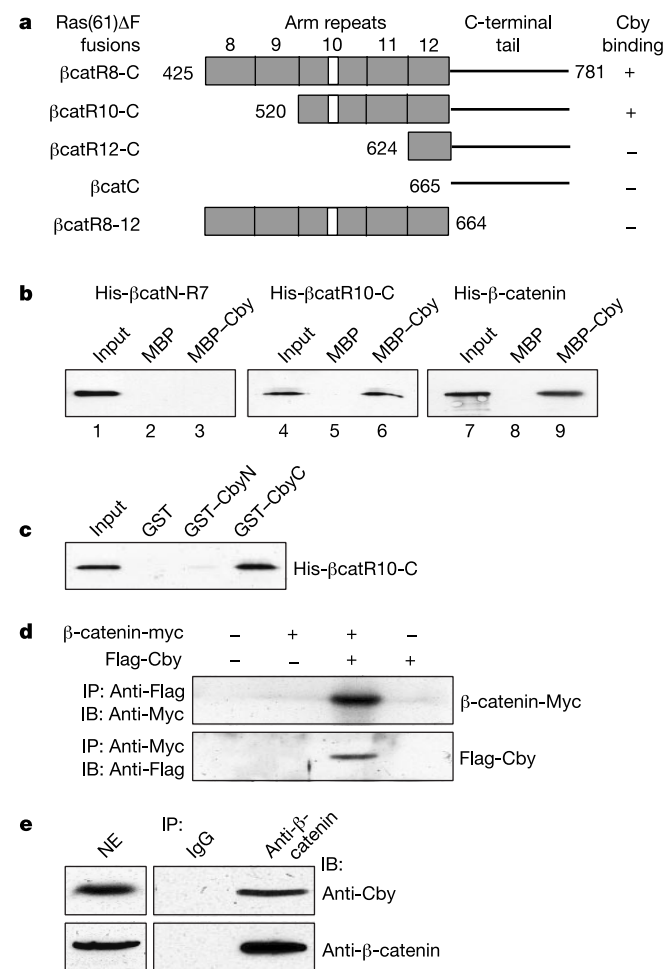


Figure 3 Physical interaction between Cby and β -catenin. **a**, The minimum domain of β -catenin required for Cby binding was mapped by RRS in yeast (see ref. 11 for details of constructs). **b**, Direct binding of Cby to β -catenin in *in vitro* binding assays. A recombinant MBP-Cby fusion protein and MBP alone were tested for their ability to pull down His-tagged β -catenin proteins as indicated, followed by Western blotting with anti-His antibodies. His- β catN-R7 and His- β catR10-C proteins contain the N-terminus to Arm repeat 7 and Arm repeat 10 to the C-terminus of β -catenin, respectively. **c**, The C-terminal half of Cby associates with β -catenin. GST fusion proteins with the N-terminal (amino acids 1–63) or C-terminal (amino acids 64–126) half of Cby were tested for their ability to pull down the His- β catR10-C protein. **d**, *In vivo* association. Expression vectors were transfected into 293T cells as shown. β -Catenin co-immunoprecipitates with Cby and vice versa. **e**, Interaction between the endogenous proteins. Nuclear extracts from 293T cells were subjected to immunoprecipitation with control IgG or an anti- β -catenin antibody and immunoblotted with the antibodies indicated. NE, nuclear extract; IP, immunoprecipitation; IB, immunoblot.

represses expression of *en*. Because *en* is in turn required to maintain *wg* expression, *wg* activates its own expression through a paracrine feedback loop¹⁹. RNAi depletion of *cby* resulted in expansion of *wg* messenger RNA and protein expression, further indicating that loss of *cby* function leads to *wg* pathway activation (Fig. 1j–m). The expanded width and intensity of the *Wg* stripes is greater than in embryos where *wg* signalling is maximally stimulated¹⁵. This suggests that in addition to *wg* signalling, *cby* acts in a regulatory process at present unknown.

Our experiments with mammalian Cby suggest that it inhibits Wnt signalling by binding to β -catenin in the nucleus and blocking its interaction with Tcf/Lef transcription factors. If *Drosophila* Cby functions in a similar manner, then loss of *cby* should not affect the abundance or localization of Arm (β -catenin). Arm localizes to

nuclei in stripes of *Wg*-responding cells at stage 9 of embryogenesis²¹. We examined Arm protein in *cby(RNAi)* embryos and observed that Arm abundance and localization were not detectably affected (Fig. 1g compared to controls, not shown).

If Cby were involved in transducing the *Wg* signal, then *Wg* would be predicted to lie upstream of Cby in an epistasis genetic pathway. To test for epistasis, we examined the expression of a *wg*-responsive *UbxB-lacZ* reporter gene in the embryonic midgut²². Expression is controlled by an enhancer that interacts with *Drosophila* Tcf and is activated in a manner dependent upon *Wg* signalling²³. *Wg* is expressed in parasegment (ps) 8, where it controls *UbxB-lacZ* expression in visceral mesoderm throughout ps7–9, as well as development of the second midgut constriction (Fig. 5a). In a *wg* null embryo, *UbxB-lacZ* expression is greatly reduced and the midgut constriction fails to form (Fig. 5c), indicating a strong dependence on *wg* function²². In *cby(RNAi)* embryos, *lacZ* expression expanded into anterior and posterior parasegments, and the second but not the first midgut constriction formed (Fig. 5b)—phenotypes reminiscent of moderate *wg* misexpression throughout the midgut²³. Thus, consistent with our observations of *cby* function in the epidermis, *cby* represses *wg*-dependent gene expression and development in visceral mesoderm. When *cby* was depleted by RNAi in *wg* null embryos, *UbxB-lacZ* expression was not blocked, and development of the second midgut constriction occurred (Fig. 5d). We note that the *wg cby(RNAi)* embryos resembled wild-type more than *cby(RNAi)* embryos; they sometimes exhibited more restricted *lacZ* expression and a first midgut constriction. As embryonic RNAi usually leads to reduced levels of gene product, this implies that the residual *cby* product represses visceral mesoderm more effectively in the absence of *wg*. These results establish that *wg* function is mediated, at least in part, through *cby*.

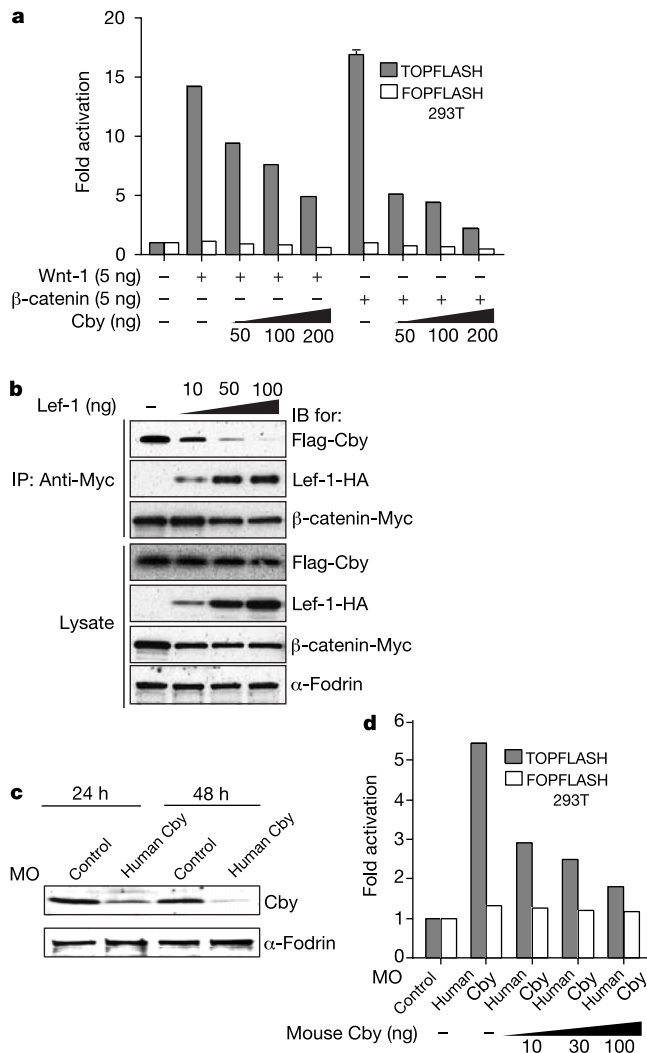


Figure 4 Cby antagonizes β -catenin signalling. **a**, 293T cells were transfected with a luciferase reporter (20 ng) and expression vectors as indicated and relative luciferase activities were measured. **b**, Cby competes with Lef-1 for binding to β -catenin. Cell lysates from 293T cells transfected with plasmids for β -catenin-Myc (150 ng), Flag-Cby (150 ng) and increasing amounts of Lef-1-HA were immunoprecipitated with an anti-Myc antibody. The immunoprecipitates or cell lysates were immunoblotted with each antibody as shown. α -Fodrin was used as a loading control. **c**, **d**, A human Cby antisense morpholino oligo (CbyMO) stimulates TOPFLASH in 293T cells. Human CbyMO decreases the endogenous protein levels 24 or 48 h after transfection compared to the control MO (**c**) and stimulates TOPFLASH in the presence of a small amount of β -catenin (1 ng), and co-transfection with a mouse Cby plasmid rescues the TOPFLASH activity to near the basal level (**d**).

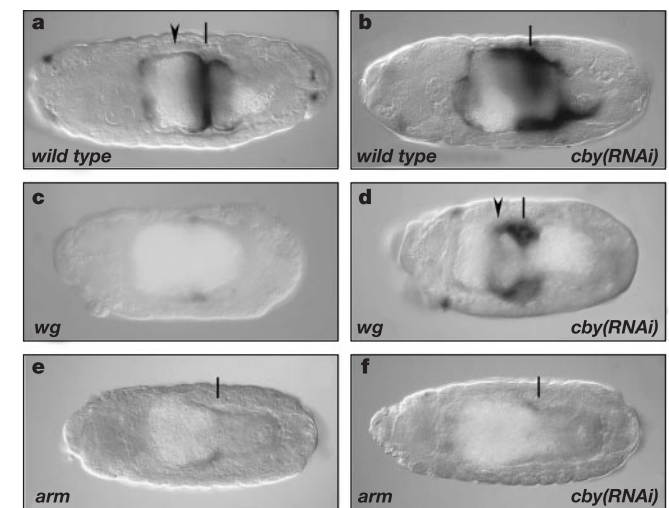


Figure 5 Genetic relationship between *cby* and genes in the *wg* pathway. Midgut expression of the *UbxB-lacZ* transgene in stage 15 embryos. The second midgut constriction is marked by vertical bars and the incipient first midgut constriction is marked by arrowheads; anterior is left and ventral down. **a**, Wild-type embryo. *UbxB-lacZ* is strongly expressed in the midgut in ps7–8 and weakly in the gastric caecum (ps3). **b**, *cby(RNAi)* results in anterior expansion of *UbxB-lacZ* expression up to ps4, and posterior expansion up to ps10. Expansion is asymmetric, a common feature of *cby(RNAi)* embryos and probably due to non-uniform interference. No first midgut constriction is evident. **c**, A *wg*¹⁻⁸ embryo. Transgene expression is repressed and no second midgut constriction has formed. **d**, A *wg*¹⁻⁸ embryo injected with *cby* dsRNA. Transgene expression is moderately expanded, and both midgut constrictions form. **e**, An *arm*^{xp33} hemizygous embryo. A vestige of a second constriction is possibly present. **f**, An *arm*^{xp33} hemizygous embryo injected with *cby* dsRNA.

We also examined the relationship between *chy* and *arm*. Embryos that are *arm chy(RNAi)* resemble *arm* embryos; they display background *lacZ* expression and strongly inhibited second midgut constriction (Fig. 5e, f). The block to midgut constriction was possibly incomplete, owing to residual *arm* activity supplied maternally (embryos depleted of maternal *arm* product exhibit a complete block to this constriction)²². The similarity of phenotypes of *arm* and *arm chy(RNAi)* embryos suggests that *chy* functions upstream of *arm* in Wg signalling, and that in wild-type embryos the role of *chy* is to repress *arm*. Our results suggest that *chy* activity is not mediated through a redistribution of β -catenin/Arm protein, but rather from inhibiting nuclear β -catenin/Arm activity. This interpretation is consistent with Cby's conserved ability to associate with β -catenin and compete with Tcf/Lef for β -catenin. Such a competitive block would account for Cby attenuating Tcf/Lef-dependent gene activation in both mammalian cells and *Drosophila* embryos.

We conclude that Cby is a nuclear protein that is conserved throughout evolution. It antagonizes Wnt/Wg signalling by inhibiting β -catenin/Arm function in mammalian cells and in *Drosophila*, raising the possibility that it may be a tumour suppressor gene. In this regard, we have found that Cby expression is significantly downregulated in thyroid and metastatic uterine tumours, and nuclear Cby staining is missing or considerably weaker in thyroid tumours (P. Swanson, K.-I.T. and R.T.M., unpublished data). Dysregulation of β -catenin signalling has been reported in these types of cancers^{24,25}, so the decreased levels of Cby expression might be relevant to tumour formation. □

Methods

RRS screening, northern analysis and Cby antibody production

RRS screening of a HeLa cell library in yeast was carried out as described¹¹. The adult human multiple-tissue northern blot (Clontech) was hybridized with a radiolabelled full-length Cby complementary DNA probe. Rabbit anti-Cby antibodies were raised against a maltose binding protein (MBP) fusion with full-length human Cby protein.

DNA constructs

The β -catenin deletion plasmids for RRS, bacterial expression vectors for His- β catR10-C and His- β -catenin, and stabilized β -catenin-Myc mammalian expression construct have been described¹¹. The luciferase reporters TOP/FOPFLASH¹² and *cyclin D1* (-1745CD1Luc)¹⁴, and Lef-1-HA plasmid¹ were as published. The Wnt-1 expression vector was a gift from M. Waterman. Bacterial expression plasmids for MBP-Cby, His- β catNR7, GST-CbyN(1–63) and GST-CbyC(64–126) were generated by standard polymerase chain reaction (PCR) cloning procedures. cDNAs encoding human and mouse Cby were expressed in pCS2+. Details of the plasmids are available upon request.

In vitro binding assays

Recombinant proteins were expressed and purified according to the manufacturer's instructions. *In vitro* binding assays were performed as described¹¹.

Immunofluorescence, co-immunoprecipitations and luciferase assays

Endogenous Cby protein was stained with anti-Cby antibodies, followed by secondary green-fluorescent Alexa anti-rabbit IgG (Molecular Probes). Nuclei were visualized with BO-PRO-3 (Molecular Probes). Transfections into 293T cells were performed using Lipofectamine Plus (Invitrogen). Co-immunoprecipitation experiments were done as described¹¹. For endogenous protein interactions, nuclear extracts were prepared from 293T cells as described²⁶ and immunoprecipitated with control goat IgG or a goat anti- β -catenin antibody (Santa Cruz). The immunoprecipitates were eluted with the competitive peptide for the anti- β -catenin antibody (Santa Cruz). The anti-cyclin A antibody is from Santa Cruz; the anti-TFIIIB antibody is from Transduction Laboratories. Luciferase assays were performed using the Dual-Luciferase Reporter Assay System (Promega). Luciferase activities were measured 24 h after transfection and normalized for transfection efficiency using the internal renilla luciferase control.

Antisense morpholino experiments

The human CbyMO complementary to the translational initiation site, 5'-GAACGTAT TCCCAAGAAAGGCATC-3', and a random control MO were synthesized by GeneTools, LLC, and were transfected by the special delivery technique as recommended by the manufacturer one day before DNA transfection.

Fly genetics and RNAi

Synthesis of *chy* dsRNA corresponding to coding sequence +3 to +273 (CG13415 in the annotated *Drosophila* genome) used previously described methods²⁷. Two 21-nucleotide DNA-RNA chimaeric oligos corresponding to *chy* coding sequence +275 to +283 were

chemically synthesized (Dharmacon): UGCUGGCCGAGCGCAGCGdTdT (sense), CCGUGCCGCUCCGCCAGCAdTdT (antisense). Oligos were annealed to form a siRNA duplex as described²⁸. dsRNA and siRNA were injected into syncytial blastoderm embryos, which were further incubated under halocarbon oil as described²⁷. To generate WIZ-*chy* transgenic lines, *chy* DNA (coding sequence +3 to +273) was cloned as an inverted repeat into pWIZ¹⁷, and the WIZ-*chy* construct was transformed into *Drosophila* by standard methods. Fly stocks used were: *wg*¹⁻⁸, *arm*-GAL4, *nos*-GAL4, UAS-*wg*¹⁸, *Ubx*-*lacZ*, and *arm*^{XP33}. For epistasis analysis, a CyO, *fz*-*lacZ* balancer was used to identify homozygous *Ubx*-*lacZ* *wg*¹⁻⁸ embryos. To identify *arm*^{XP33} hemizygous embryos laid from *arm*^{XP33}/FM7c females, they were incubated with rabbit anti- β -galactosidase and mouse anti-Arm antibody. Embryos were treated with Alexa543 anti-mouse secondary, and those that did not exhibit Arm fluorescence were identified.

Cuticle analysis, immunohistochemistry and in situ hybridization

For cuticle preparations, embryos were incubated under oil and their cuticles were prepared as described²⁷. For X-gal staining and immunohistochemistry, embryos were incubated at 18 °C to the appropriate stage. They were fixed, devitellinized and stained as described²⁷. Primary antibodies were rabbit anti- β -galactosidase (Cappel), mouse anti-Arm (Mab N27A1), mouse anti-En (Mab 4D9), or rabbit anti-Wg antibody. Secondary antibodies used were Alexa-conjugated IgGs (Molecular Probes) or HRP-conjugated IgGs (Bio-Rad). For *in situ* hybridization, embryos were fixed in heptane:3.7% formaldehyde in PEM (1:1) for 15 min, equilibrated in PBS, and dissected out of their vitelline membranes. Embryos were washed in PBW (PBS + 0.1% Tween20), 50% methanol in PBW, 100% methanol, 50% methanol in PBW, and four changes of PBW. They were fixed for 20 min in 5% paraformaldehyde, 25 mM EGTA, 0.05% Tween20 in PBS, digested in 6 μ g ml⁻¹ proteinase K for two minutes, blocked in glycine buffer, and post-fixed for 20 min in 5% paraformaldehyde, 25 mM EGTA, 0.05% Tween20 in PBS. Embryos were washed thoroughly in PBW, followed by a wash in 50% PBW: 50% hybridization (Hyb) buffer. They were then denatured for 10 min at 75 °C in formamide:TE (1:1), followed by plunging in ice. Embryos were hybridized to probe in Hyb buffer, washed and stained as described²⁹.

Received 19 December 2002; accepted 11 March 2003; doi:10.1038/nature01570.

- Miller, J. R., Hocking, A. M., Brown, J. D. & Moon, R. T. Mechanism and function of signal transduction by the Wnt/ β -catenin and Wnt/Ca²⁺ pathways. *Oncogene* **18**, 7860–7872 (1999).
- Polakis, P. Wnt signaling and cancer. *Genes Dev.* **14**, 1837–1851 (2000).
- Behrens, J. et al. Functional interaction of β -catenin with the transcription factor Lef-1. *Nature* **382**, 638–642 (1996).
- Molenaar, M. et al. XTcf-3 transcription factor mediates β -catenin-induced axis formation in *Xenopus* embryos. *Cell* **86**, 391–399 (1996).
- Brunner, E., Peter, O., Schweizer, L. & Basler, K. *pangolin* encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in *Drosophila*. *Nature* **385**, 829–833 (1997).
- van de Wetering, M. et al. Armadillo coactivates transcription driven by the product of the *Drosophila* segment polarity gene *TCF*. *Cell* **88**, 789–799 (1997).
- Peifer, M., Berg, S. & Reynolds, A. B. A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell* **76**, 789–791 (1994).
- Hsu, S. C., Galceran, J. & Grosschedl, R. Modulation of transcriptional regulation by Lef-1 in response to Wnt-1 signaling and association with β -catenin. *Mol. Cell Biol.* **18**, 4807–4818 (1998).
- Hecht, A., Litterst, C. M., Huber, O. & Kemler, R. Functional characterization of multiple transactivating elements in β -catenin, some of which interact with the TATA-binding protein *in vitro*. *J. Biol. Chem.* **274**, 18017–18025 (1999).
- Broder, Y. C., Katz, S. & Aronheim, A. The Ras recruitment system, a novel approach to the study of protein-protein interactions. *Curr. Biol.* **8**, 1121–1124 (1998).
- Takemaru, K.-I. & Moon, R. T. The transcriptional coactivator CBP interacts with β -catenin to activate gene expression. *J. Cell Biol.* **149**, 249–254 (2000).
- Korinek, V. et al. Constitutive transcriptional activation by a β -catenin-Tcf complex in APC^{-/-} colon carcinoma. *Science* **275**, 1784–1787 (1997).
- Yost, C. et al. The axis-inducing activity, stability, and subcellular distribution of β -catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3. *Genes Dev.* **10**, 1443–1454 (1996).
- Shtutman, M. et al. The cyclin D1 gene is a target of the β -catenin/LEF-1 pathway. *Proc. Natl Acad. Sci. USA* **96**, 5522–5527 (1999).
- Noordermeer, J., Johnston, P., Rijsewijk, F., Nusse, R. & Lawrence, P. A. The consequences of ubiquitous expression of the *wingless* gene in the *Drosophila* embryo. *Development* **116**, 711–719 (1992).
- Peifer, M. & Wieschaus, E. The segment polarity gene *armadillo* encodes a functionally modular protein that is the *Drosophila* homolog of human plakoglobin. *Cell* **63**, 1167–1176 (1990).
- Lee, Y. S. & Carthew, R. W. Making a better RNAi vector for *Drosophila*: Use of intron spacers. *Methods* (in the press).
- Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
- Cadigan, K. M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- Sanson, B., Alexandre, C., Fascetti, N. & Vincent, J.-P. Engrailed and hedgehog make the range of wingless asymmetric in *Drosophila* embryos. *Cell* **98**, 207–216 (1999).
- Peifer, M., Sweeten, D., Casey, M. & Wieschaus, E. *wingless* signal and Zeste-white 3 kinase trigger opposing changes in the intracellular distribution of Armadillo. *Development* **120**, 369–380 (1994).
- Yu, X., Hoppler, S., Eresh, S. & Bienz, M. *decapentaplegic*, a target gene of the *wingless* signalling pathway in the *Drosophila* midgut. *Development* **122**, 849–858 (1996).
- Yu, X., Riese, J., Eresh, S. & Bienz, M. Transcriptional repression due to high levels of Wingless signalling. *EMBO J.* **17**, 7021–7032 (1998).
- Garcia-Rostan, G. et al. β -Catenin dysregulation in thyroid neoplasms. *Am. J. Pathol.* **158**, 987–996 (2001).

25. Fukuchi, T. *et al.* β -Catenin mutation in carcinoma of the uterine endometrium. *Cancer Res.* **58**, 3526–3528 (1998).
26. Andrews, N. C. & Faller, D. V. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19**, 2499 (1991).
27. Kennerdell, J. R. & Carthew, R. W. Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled2* act in the *wingless* pathway. *Cell* **95**, 1017–1026 (1998).
28. Carthew, R. W. in *Gene Silencing* (ed. Hannon, G.) (Cold Spring Harbor Laboratory Press, New York, 2003).
29. Tautz, D. & Pfeifle, C. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81–85 (1988).

Supplementary Information accompanies the paper on Nature's website
 (http://www.nature.com/nature).

Acknowledgements We thank R. Holmgren, R. Pestell, M. Waterman and DSHB for reagents; M. Peifer, J. Treisman and Bloomington Stock Center for fly stocks; R. Holmgren, E.-Q. Li and members of the Carthew and Moon laboratories for critically reading the manuscript. This work was supported by postdoctoral fellowships from the Japan Science and Technology Corporation and Uehara Memorial Foundation (to K.-I.T.), the Pew Charitable Trust (to R.W.C.), and by an NIH grant (to R.T.M.). R.T.M. is an Investigator of the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.T.M.
 (e-mail: rtmoon@u.washington.edu). GenBank accession numbers for Cby are: human, AL050345; mouse, AK003719; *Xenopus*, BJ093998; zebrafish, BI885798; *Drosophila*, AE003736.

DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing

Alexander Saveliy^{*,†}, Christopher Everett^{*,†}, Tammy Sharpe^{*,}
 Zoë Webster[‡] & Richard Festenstein^{*,§}

^{*} CSC Gene Control Mechanisms and Disease Group, Experimental Genetics, Faculty of Medicine, Imperial College School of Medicine, Hammersmith Campus, Du Cane Road, London W12 0NN, UK

[‡] Transgenics and Embryonic Stem Cell Laboratory, MRC Clinical Sciences Centre, Hammersmith Campus, Du Cane Road, London W12 0NN, UK

[§] Institute of Neurology Department of Neurogenetics, National Hospital for Neurology and Neurosurgery, Queen Square, London WC1N 3BG, UK

[†] These authors contributed equally to this work.

Gene repression is crucial to the maintenance of differentiated cell types in multicellular organisms, whereas aberrant silencing can lead to disease. The organization of DNA into chromatin and heterochromatin¹ is implicated in gene silencing. In chromatin, DNA wraps around histones, creating nucleosomes. Further condensation of chromatin, associated with large blocks of repetitive DNA sequences, is known as heterochromatin. Position effect variegation (PEV) occurs when a gene is located abnormally close to heterochromatin, silencing the affected gene in a proportion of cells¹. Here we show that the relatively short triplet-repeat expansions found in myotonic dystrophy and Friedreich's ataxia confer variegation of expression on a linked transgene in mice. Silencing was correlated with a decrease in promoter accessibility and was enhanced by the classical PEV modifier heterochromatin protein 1 (HP1). Notably, triplet-repeat-associated variegation was not restricted to classical heterochromatic regions but occurred irrespective of chromosomal location. Because the phenomenon described here shares important features with PEV, the mechanisms underlying heterochromatin-mediated silencing might have a role in gene regulation at many sites throughout the mammalian genome and

modulate the extent of gene silencing and hence severity in several triplet-repeat diseases.

Although recent studies on modifiers of PEV have identified molecules (such as SUV39 and HP1) that participate to propagate heterochromatic structures^{2,3}, very little is known about the DNA *cis*-acting requirements for heterochromatin formation. Certain transgenes are prone to variegation when repeated in tandem arrays^{4,5}; however, some transgenes do not variegate despite being repeated in tandem¹. Gene regulatory elements known as locus control regions (LCRs)⁶ can overcome pericentromeric PEV and disabling them renders the expression of linked genes PEV-sensitive^{7,8}. Aberrant gene silencing can occur in diseases—for example, deletion of an LCR results in thalassaemia⁹—whereas the expansion of GAA or CTG triplet repeats leads to gene silencing in Friedreich's ataxia¹⁰ and myotonic dystrophy^{11,12}, respectively. Consistent with the regular spacing of nucleosomes found in heterochromatin

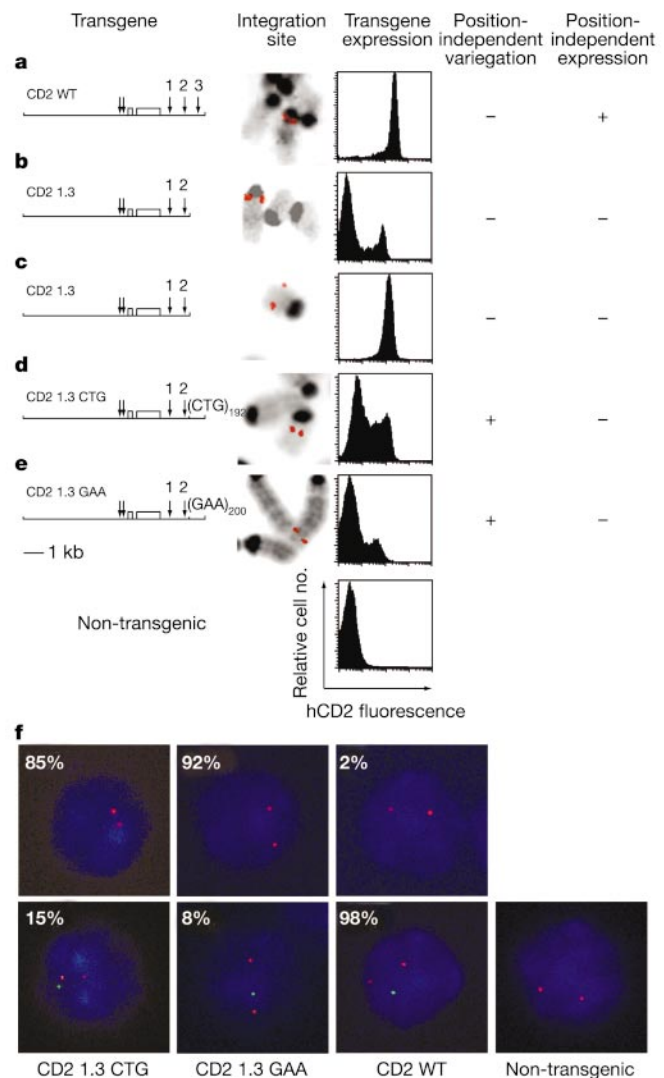


Figure 1 CTG or GAA triplet-repeat expansions confer position-independent variegation on a heterochromatin-sensitive reporter. **a**, CD2 wild type, the hCD2 minigene with LCR⁷, expressed hCD2 on all T cells despite a pericentromeric location. **b**, **c**, Transgenes lacking LCR HSS3 (CD2 1.3) showed variegation when pericentromeric (**b**) and unimodal expression when in the chromosomal long arm (**c**). **d**, **e**, However, when linked to 192 CTG (**d**) or 200 GAA (**e**) repeats, expression of the CD2 1.3 transgene was variegated irrespective of location. kb, kilobase. **f**, Primary-transcript RNA FISH. Hybridization of thymocytes from CD2 1.3 CTG and CD2 1.3 GAA lines with mouse CD2 (red) and hCD2 (green) probes revealed hCD2 variegation at the transcriptional level.

Table 1 CTG and GAA repeats confer variegation on a linked hCD2 transgene irrespective of its chromosomal location

	CD2 with full LCR†	CD2 1.3‡	CD2 1.3 CTG	CD2 1.3 GAA
Pericentromeric integration	U (n = 3)	V (n = 4)	V (n = 5)	V (n = 1)
Peritelomeric integration	– (n = 0)	U (n = 1) V (n = 1)	V (n = 2)	V (n = 2)
Integration in long arm of chromosome	U (n = 5)	U (n = 4) V (n = 2)	V* (n = 3)	V* (n = 5)
Variegation	None	Position-dependent variegation	Position-independent variegation	Position-independent variegation

V, variegates; U, unimodal expression; n, number of transgenic lines.

* $P < 0.05$, exact one-sided midpoint P values. Contingency tables were constructed to compare numbers of CD2 1.3 variegating lines (with chromosomal long-arm transgene integration sites) that showed variegation with and without triplet repeats. The STATXACT program (Cytel Corporation, Massachusetts, USA) was used to analyse the results and obtain exact probabilities.

†The figures shown for these transgenic lines were obtained in a previous study⁷.

‡These figures combine the results obtained from six lines analysed previously⁷ with the results obtained after an identical analysis of six new lines.

*in vivo*¹³ is the observation that CTG repeats 'position' nucleosomes *in vitro*¹⁴. Here we investigate the hypothesis that such triplet repeat expansions might act to stimulate heterochromatin formation. In myotonic dystrophy, a CTG expansion in the untranslated part of the dystrophin myotonia protein kinase (*DMPK*) gene, which is located immediately upstream of the *Six 5* gene, results in *Six 5* downregulation and decreased DNase I accessibility at its promoter¹⁵. Mice deficient in *Six 5* develop cataracts—a feature of myotonic dystrophy¹⁶. Although *DMPK* might also be downregulated in patients with myotonic dystrophy^{17,18}, the CUG-containing RNA forms nuclear aggregates¹⁹ and has a toxic gain-of-function effect resulting in the disease phenotype²⁰. In Friedreich's ataxia, GAA expansion in the intron of the *X25* gene causes its downregulation. Although studies *in vitro* suggest that expanded GAA repeats can form a triplex DNA structure²¹, thereby impeding transcriptional elongation²², their effect on chromatin packaging is unknown.

To determine whether CTG or GAA expansions could silence genes at the level of chromatin packaging, we used a heterochromatin-sensitive human CD2 (hCD2) transgene as a reporter. There are three DNase I-hypersensitive sites (HSSs) within the hCD2 LCR (Fig. 1a), which directs hCD2 expression to all T cells in transgenic mice irrespective of the chromosomal location of the transgene^{7,23}. The omission of HSS3 (Fig. 1b and c) renders the transgene (CD2 1.3) susceptible to silencing by heterochromatin (for example by pericentromeric integrations) resulting in variegation of hCD2 expression (Fig. 1b). In contrast, CD2 1.3 transgenes, which integrate in euchromatic locations (such as the chromosomal long arm; Fig. 1c), express hCD2 on all T cells⁷ over a range (4–25 copies) of transgene copy numbers⁷. By generating and analysing (as described previously⁷) six additional CD2 1.3 transgenic lines, the data presented here extend the initial observation that the CD2 1.3 transgene is sensitive to chromosomal location (see Table 1) and that large CD2 1.3 tandem arrays (up to 25 copies) are not sufficient to induce variegation. The variegation observed in some non-centromeric locations suggests that this reporter can be used as a 'gene trap' to identify heterochromatic sites in the genome that cannot be identified cytogenetically²⁴. These properties of the CD2 1.3 transgene allowed us to assay the heterochromatin-forming potential of DNA sequences *in vivo*.

We reasoned that if expanded triplet repeats were capable of mediating a heterochromatin-like effect on a juxtaposed CD2 1.3 transgene, the outcome would be variegation of hCD2 expression irrespective of the chromosomal location of the transgenes. We linked 192 CTG or 200 GAA repeats to the CD2 1.3 reporter gene (Fig. 1d and e) to generate ten CD2 1.3 CTG and nine CD2 1.3 GAA transgenic mouse lines. Southern blot analysis revealed each line to contain 'head-to-tail' (5' → 3') tandem arrays with a range of transgene copy numbers from 3 to 25 for the CD2 1.3 CTG lines and from 3 to 27 for the CD2 1.3 GAA lines. Each line contained about 190 CTG repeats or 200 GAA repeats as well as some expanded and contracted repeats that remained in the pathological range (namely 70–350 CTG repeats, or 40–500 GAA repeats); somatic repeat instability was undetectable at 4–8 weeks of age.

Transgene location was determined by fluorescence *in situ* hybridization (FISH) and hCD2 expression by flow cytometry using fluorescently labelled anti-hCD2 antibody⁷. The mean fluorescence detected on the surface of the T cells was previously shown to be directly correlated with the steady-state hCD2 messenger RNA levels (see Supplementary Fig. S1a). Strikingly, expression of hCD2 on thymocytes and mature T cells was variegated in all lines and occurred at pericentromeric, peritelomeric and long-arm locations, and was therefore not dependent on chromosomal position (see Table 1). Three CD2 1.3 CTG lines and five CD2 1.3 GAA lines that had their transgenes located at different sites in the chromosomal long arm all showed hCD2 variegation (this was statistically significant at $P < 0.05$ when compared with CD2 1.3 lines lacking the repeats with long-arm integration sites; see Table 1). Examples of this are shown in Fig. 1d and e. The histograms demonstrate bimodal hCD2 expression on T cells typical of variegation. That the variegation in these lines was occurring at the level of transcription was shown by using primary-transcript RNA FISH on cells from two independent CD2 1.3 CTG and CD2 1.3 GAA lines (examples are shown in Fig. 1f); northern blot analysis from these lines confirmed marked transcriptional repression (Supplementary

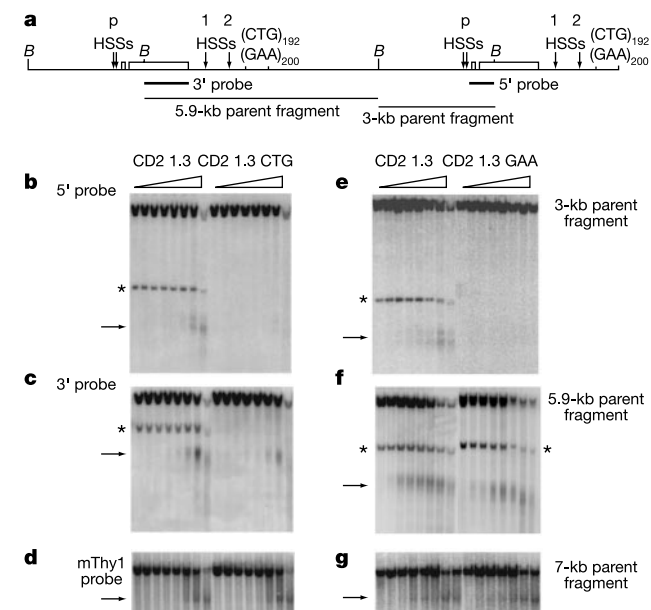


Figure 2 Triplet-repeat-associated gene silencing correlated with inaccessible chromatin at the promoter but not the enhancer. **a**, Transgene map. **B**, *Bgl*II. **b–g**, DNase I hypersensitivity analysis⁷ on thymocyte nuclei from CD2 1.3 CTG transgenics (**b**, **c**, **d**) or CD2 1.3 GAA transgenics (**e**, **f**, **g**) compared with CD2 1.3 nuclei lacking the triplet repeats. Transgenes were all in the chromosomal long arm. **b**, **e**, Promoter (p) HSS analysis; **c**, **f**, enhancer HSS analysis; **d**, **g**, *Thy1* HSS analysis as an internal control⁷. Arrows indicate bands resulting from treatment with DNase I. Asterisk indicates transgene end-fragments, the length of which was determined by mouse genome *Bgl*II sites.

Fig. S1b). Southern blot analysis on DNA obtained from T cells sorted according to hCD2 expression excluded the possibility that this bimodal hCD2 expression pattern resulted either from deletion of the transgene in the cells not expressing hCD2 or from truncation of triplet repeats in the hCD2-expressing cells (data not shown). Thus, the inclusion of CTG or GAA repeats within the CD2 1.3 transgene leads to chromosomal position independent variegation.

To investigate whether CTG or GAA repeat expansions could be mediating their effect by modifying chromatin structure, DNase I (Fig. 2) and micrococcal nuclease (MNase) (Fig. 3) accessibility studies were undertaken on transgenics with chromosomal long arm transgene locations. The former, to examine the DNase I hypersensitive sites, and the latter to assess the arrangement of nucleosomes proximal to the repeats. Aliquots of nuclei from transgenic thymocytes (of which less than 10% expressed the hCD2 transgene) were exposed to increasing concentrations of nuclease and the DNA obtained analysed by Southern blot as previously described⁷. Thymocyte nuclei from hCD2 transgenic mice (CD2 1.3), which express hCD2 on all T cells, were used as

controls (first 8 lanes in each Southern blot – Fig. 2b–g). The top blots (Fig. 2b and e) were hybridized with the 5' probe (Fig. 2a), thereby identifying the promoter HSS in CD2 1.3 transgenic thymocytes. Strikingly, this HSS was undetectable in nuclei obtained from either the CD2 1.3 CTG or the CD2 1.3 GAA transgenic mice (Fig. 2b and e). However, hybridization with the 3' probe revealed that the HSS located within the truncated LCR was formed irrespective of whether the CTG or GAA repeats were present (Fig. 2c and f). (The analysis represented in Fig. 2 was reproduced for an additional CD2 1.3 CTG and CD2 1.3 GAA line.) Thus, these repeats appear to specifically inhibit the formation of the promoter HSS, leaving the LCR HSS1 (previously shown to be an enhancer²⁵) relatively unaffected. In contrast, both the promoter and enhancer HSS were undetectable when a CD2 1.3 transgene was silenced by proximity to pericentromeric heterochromatin^{7,24}. This result may reflect differences in the 'chromatin packaging' ability of different repetitive DNA sequences (triplet repeats vs pericentromeric repeats) and is consistent with a probability/competition model for chromatin packaging⁹ in which the enhancer is more

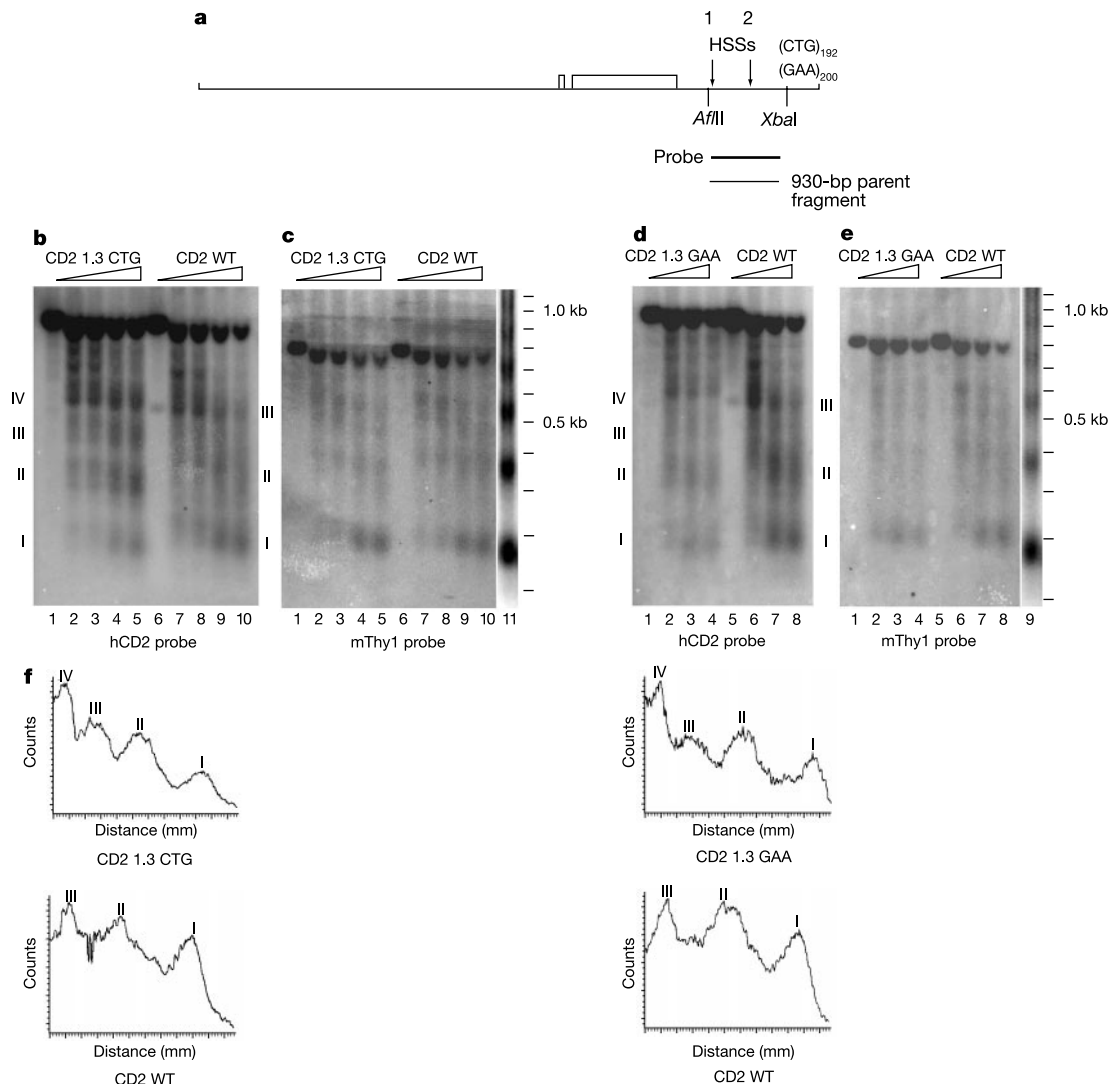


Figure 3 Increased nucleosomal density in triplet-repeat proximal regions. **a**, DNA from MNase-treated thymocyte nuclei was digested to release a 930-bp fragment adjacent to the triplet-repeat insertion, and nucleosomal density was estimated by Southern blotting. WT, wild type. **b**, A more distinct and compact nucleosomal ladder is found adjacent to the CTG repeat (lanes 1–5) than in its absence (lanes 6–10). **c**, The same DNA probed for

endogenous *Thy1*. **d**, DNA adjacent to the GAA repeat had a similar pattern to that in **b**, **e**. The same DNA analysed with the *Thy1* probe. The average nucleosomal density is shown by ethidium bromide staining at the far right. **f**, Phosphorimager quantification of lines (origin 0.6 kb) drawn downwards through lanes 4 and 9 in **b** and 3 and 7 in **d**.

effective than the promoter at competing with silencing factors to establish and maintain an open chromatin configuration. In addition, hypersensitivity at the promoter may be dependent on its interaction with the LCR hypersensitive site. The abnormal chromatin packaging induced by triplet repeats may inhibit this interaction. We therefore examined the chromatin structure of the region immediately adjacent to the repeats which contains the LCR HSS. Consistent with the finding that CTG repeats are very efficiently packaged into nucleosomal arrays *in vitro*¹⁴, our MNase studies indicated that, in the presence of the triplet repeats, the adjacent DNA had a more dense and ordered nucleosomal array than the wild type hCD2 transgene (Fig. 3b and f). More distinct bands were obtained with less background signal in the region proximal to the CTG repeats (Fig. 3b) compared to the same region when adjacent to the GAA repeats (Fig. 3d). This suggests that CTG repeats induce more precise nucleosome positioning than GAA repeats. The endogenous *Thyl* gene (which is transcribed in T cells and was used as a control for the extent of MNase digestion) showed a lower density of nucleosomal packaging (Fig. 3c and e) similar to the hCD2 wild type transgene which contained no triplet repeats. Notably, a recent study in *Drosophila* used the same approach to demonstrate that a transgene located in constitutive heterochromatin had a more compacted nucleosomal structure compared with an identical transgene located in euchromatin¹³. Thus, the inclusion of CTG or GAA repeats within the CD2 1.3 transgene not only leads to position-independent variegation but also correlates with a 'closed' chromatin structure at the promoter of the transgene. It may be that the promoter hypersensitive site cannot form because the triplet-repeat-associated heterochromatin inhibits the interaction of the enhancer hypersensitive site with the promoter. Such direct 'looping' may be essential for activation of the promoter and

has recently been shown to occur between the LCR and the β -globin locus²⁶.

Recently, investigation of PEV modifiers has revealed mechanisms involved in heterochromatin formation. A powerful modifier of PEV is the methyltransferase Suv3-9 that establishes the binding site for HP1 by methylating Lys 9 on the tail of histone H3. Suv3-9 and HP1 are thought to associate with each other, thereby propagating heterochromatin along the chromosome^{2,3}. Most PEV studies use genes located near cytogenetically defined heterochromatin such as the centromere. A key question is therefore whether similar heterochromatin-mediated silencing can occur elsewhere in the mammalian genome. We had previously shown that overexpression of the classical PEV modifier HP1 β enhanced the silencing of pericentromeric CD2 1.3 transgenes in transgenic mice²⁴. Because transgenes containing CTG or GAA repeats *in cis* exhibit variegation of expression even when located in the chromosomal long arm, we investigated whether HP1 β dosage could influence the extent of this silencing. CD2 1.3 CTG and CD2 1.3 GAA transgenic mouse lines that carried their transgenes in the chromosomal long arm were crossed with transgenic mice overexpressing HP1 β (fivefold increase over endogenous concentrations at the protein level; data not shown). Analysis of the offspring showed that HP1 β overexpression, in both CD2 1.3 CTG and CD2 1.3 GAA transgenic mouse lines, significantly decreased the percentage of hCD2-expressing T cells (Fig. 4a and b), whereas in the absence of triplet repeats and variegation, the overexpression of HP1 β had no effect on the proportion of expressing cells²⁴ (Fig. 4c and d). This result further implicates heterochromatin formation as the mechanism underlying this GAA or CTG repeat-mediated gene silencing.

Thus, CTG and GAA trinucleotide repeats, which expand pathologically in human disease, might exert their gene-silencing effect by a mechanism resembling heterochromatin-mediated PEV. Although it remains to be seen whether other examples of the many repeats found in vertebrate genomes are capable of similar effects, it is tempting to speculate that the opposing effects of LCRs and repetitive DNA on chromatin structure and gene expression are a reflection of their evolutionary interdependence. That HP1 has a role in gene repression is suggested by two recent studies: one correlated the binding of HP1 to retinoblastoma protein with the silencing of the cyclin E gene²⁷, and the other correlated methylation of H3 Lys 9 with gene silencing²⁸. A recent study in *Drosophila* suggested that HP1 participated in gene repression at two euchromatic locations²⁹. To our knowledge the findings presented here provide the first evidence that PEV-like silencing can occur at multiple sites in the mammalian chromosomal long arm where most genes are located and therefore indicate that the modifiers of pericentromeric PEV might participate in physiological gene silencing in mammals. Our findings show that heterochromatin-like effects can be mediated by short CTG and GAA triplet repeats. Furthermore, because the repeats studied here cause gene silencing in human disease it is possible that the modifiers of PEV might modify clinical severity. □

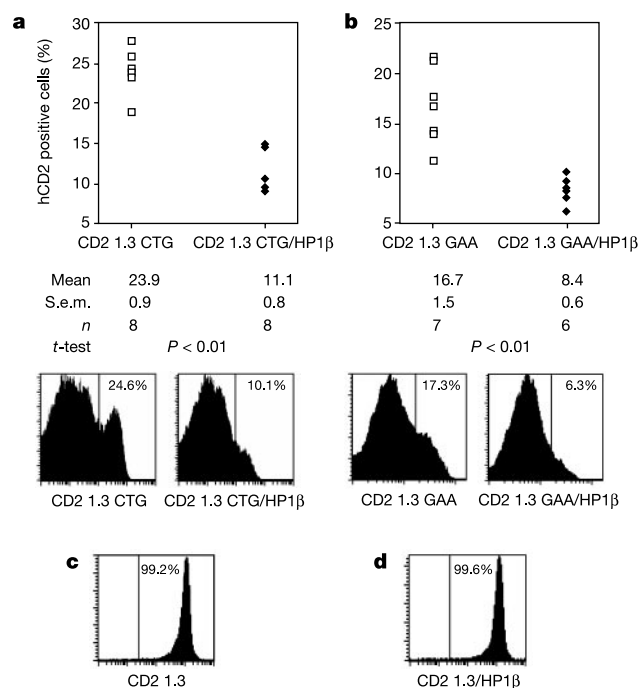


Figure 4 Triplet-repeat-associated variegation is enhanced by overexpression of HP1 β . **a, b**, Scatter plots showing proportions of hCD2-expressing T cells. Representative histogram plots are shown below. **a**, CD2 1.3 CTG mice were compared with siblings overexpressing HP1 β . **b**, A similar analysis of CD2 1.3 GAA mice. The result was reproduced in two independent CD2 1.3 CTG and CD2 1.3 GAA lines. **c, d**, Overexpressing HP1 β (**d**) in a CD2 1.3 line (i.e. lacking triplet repeats) that does not variegate had no effect on the proportion of hCD2-expressing T cells in comparison with the control (**c**).

Methods

Mice

All recombinant DNA manipulations were performed with standard techniques. The plasmids containing the expanded CTG and GAA sequences were a gift from Mauro Santibanez-Koref. The CTG or GAA repeat was subcloned at the 3' end of the hCD2 1.3 construct¹² by using the *KpnI*-*XbaI* or *KpnI*-*BamHI* sites, respectively. Sequencing confirmed that about 190 CTG repeats and 200 GAA repeats were contained within the plasmids and the joints were intact. The resulting hCD2 1.3 CTG and hCD2 1.3 GAA constructs were lifted out with the *SalI* and *NorI* enzymes, purified with Elutip-D columns (Schleicher & Schuell) and injected into mouse C57Black6/CBA Ca fertilized eggs. Founder transgenic mice were crossed with CBA Ca mice to establish transgenic lines. The project was given ethical approval by Imperial College and the UK Home Office.

Flow cytometric and FISH analysis

Flow cytometric and FISH analyses of second-generation transgenic mice were performed as described previously⁷. Primary-transcript RNA FISH was done essentially as described

previously³⁰. In brief, thymocyte suspensions were prepared and left to settle for 90 s on poly(L-lysine)-coated slides (Sigma), which were then fixed in 4% formaldehyde, 5% acetic acid in normal saline for 20 min at 23 °C. This was followed by three sequential 5-min washes in PBS. Slides were then equilibrated with 70% ethanol for 5 min followed by 5 min in 0.1 M Tris-HCl pH 7.5, 0.15 M NaCl. This was followed by digestion in 0.01% pepsin in 0.01 M HCl for 5 min at 37 °C, a brief rinse in water and fixation in 3.7% formaldehyde in PBS at 23 °C. They were then washed and dehydrated sequentially in 70%, 90% and 100% ethanol (3 min in each) and air-dried. Single-stranded biotinylated mouse CD2 intron probe (originally cloned after PCR of a mouse CD2 cosmid—a gift from D. Kioussis—using the primer pair 5′-GTTACTCCACCCCTCTTCAA-3′ and 5′-GGCACTGTTTTCCTGTTACTC-3′) and a digoxigenin-labelled human CD2 cDNA probe were prepared by *in vitro* transcription followed by reverse transcription. The probes were denatured by heating at 80 °C for 5 min. Hybridization overnight in a humidified chamber was followed by sequential washes in 2 × SSC, then by blocking and detection with several layers of antibodies. For digoxigenin the antibodies were sheep anti-digoxigenin (Roche), rabbit anti-(sheep fluorescein isothiocyanate (FITC)) (Calbiochem) and goat anti-(rabbit FITC); for the biotinylated probe the antibodies were avidin D–Texas Red (Vector), goat anti-(avidin D–biotin) (Vector) and a final step with avidin D–Texas Red. Both intron and exon probes have been shown to identify primary transcripts generated at the site of transcription³⁰. This procedure strips most of the cytoplasm from the cell, and leaves intact nuclei that were detected by staining with 4,6-diamidino-2-phenylindole. For microscopy, images were acquired with a Deltavision system (Applied Precision) and a cooled charge-coupled device camera (Photometrics CH350L) on an Olympus IX70 inverted microscope, with a 100 × 1.35 numerical aperture UPlanApo objective lens.

Nuclease sensitivity

Aliquots of nuclei obtained from 10⁸ cells²⁴ were digested either for 4 min at 37 °C with increasing amounts of DNase I (0–6 µg ml^{−1}) or for 5 min at 23 °C with MNase (Roche 107921; 0–500 U ml^{−1}). After inactivation of nuclease, nuclei were treated with proteinase K and DNA was purified, digested and analysed by Southern blotting. The *Thy1* probe used for MNase analysis was a 749-base-pair *Pst*I–*Bam*HI fragment from the fourth exon of the endogenous *Thy1* gene.

Received 30 December 2002; accepted 12 March 2003; doi:10.1038/nature01596.

- Dillon, N. & Festenstein, R. Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet.* **18**, 252–258 (2002).
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
- Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromatin domain. *Nature* **410**, 120–124 (2001).
- Dorer, D. R. & Henikoff, S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**, 993–1002 (1994).
- Garrick, D., Fiering, S., Martin, D. I. & Whitelaw, E. Repeat-induced gene silencing in mammals. *Nature Genet.* **18**, 56–59 (1998).
- Grosfeld, F., van Assendelft, G. B., Graeves, D. R. & Kollias, G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975–985 (1987).
- Festenstein, R. *et al.* Locus control region function and heterochromatin-induced position effect variegation. *Science* **271**, 1123–1125 (1996).
- Milot, E. *et al.* Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* **87**, 105–114 (1996).
- Kioussis, D. & Festenstein, R. Locus control regions: overcoming heterochromatin-induced gene inactivation in mammals. *Curr. Opin. Genet. Dev.* **7**, 614–619 (1997).
- Campuzano, V. *et al.* Friedrich's ataxia: Autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271**, 1423–1427 (1996).
- Brook, J. D. *et al.* Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 (1992).
- Mahadevan, M. *et al.* Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science* **255**, 1253–1255 (1992).
- Sun, F. L., Cuaycong, M. H. & Elgin, S. C. Long-range nucleosome ordering is associated with gene silencing in *Drosophila melanogaster* pericentric heterochromatin. *Mol. Cell. Biol.* **21**, 2867–2879 (2001).
- Wang, Y., Amirhaeri, S., Kang, S., Wells, R. D. & Griffith, J. D. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* **265**, 669–671 (1994).
- Klesert, T. R., Otten, A. D., Bird, T. D. & Tapscott, S. J. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nature Genet.* **16**, 402–406 (1997).
- Klesert, T. R. *et al.* Mice deficient in Six5 develop cataracts: Implications for myotonic dystrophy. *Nature Genet.* **25**, 105–109 (2000).
- Carango, P., Noble, J. E., Marks, H. G. & Funanage, V. L. Absence of myotonic dystrophy protein kinase (DMPK) mRNA as a result of a triplet repeat expansion in myotonic dystrophy. *Genomics* **18**, 340–348 (1993).
- Krahe, R. *et al.* Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. *Genomics* **28**, 1–14 (1995).
- Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H. & Housman, D. E. Expansion of a CUG trinucleotide repeat in the 3′ untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl Acad. Sci. USA* **94**, 7388–7393 (1997).
- Mankodi, A. *et al.* Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* **289**, 1769–1773 (2000).
- Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. & Wohlrab, F. The chemistry and biology of unusual DNA structures adopted by oligopurine.oligopyrimidine sequences. *FASEB J.* **2**, 2939–2949 (1988).
- Sakamoto, N., Ohshima, K., Montermini, L., Pandolfo, M. & Wells, R. D. Sticky DNA, a self-associated complex formed at long GAA.TTC repeats in intron 1 of the Frataxin gene, inhibits transcription. *J. Biol. Chem.* **276**, 27171–27177 (2001).
- Graeves, D. R., Wilson, F., Lang, G. & Kioussis, D. Human CD2 3′-flanking sequences confer high-level T-cell specific position independent gene expression in transgenic mice. *Cell* **56**, 979–986 (1989).
- Festenstein, R. *et al.* Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nature Genet.* **23**, 457–461 (1999).
- Lake, R. A., Wotton, D. & Owen, M. J. A 3′ transcriptional enhancer regulates tissue-specific expression of the human CD2 gene. *EMBO J.* **9**, 3129–3136 (1990).
- Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. & Fraser, P. Long-range chromatin regulatory interactions *in vivo*. *Nature Genet.* **32**, 623–626 (2002).
- Nielsen, S. J. *et al.* Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**, 561–565 (2001).
- Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D. & Felsenfeld, G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293**, 2453–2455 (2001).
- Hwang, K. K., Eissenberg, J. C. & Worman, H. J. Transcriptional repression of euchromatic genes by *Drosophila* heterochromatin protein 1 and histone modifiers. *Proc. Natl Acad. Sci. USA* **98**, 11423–11427 (2001).
- Wijgerde, M., Grosfeld, F. & Fraser, P. Transcription complex stability and chromatin dynamics *in vivo*. *Nature* **377**, 209–213 (1995).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank P. Fraser and especially C. Osborne (Babraham Institute) for help with the RNA FISH; S. Sharghi Namini, S. Uribe Lewis and D. Kazazi for technical assistance; M. Walport, N. Dillon, L. Aragorn and D. Kioussis for critical reading of an earlier version of the manuscript; and M. McCarthy for advice on statistical analysis. A.S. was a Wellcome Travelling Fellow and C.E. a Wellcome Clinical Fellow. This work was supported by the Muscular Dystrophy Campaign (UK), the MRC UK and a EU Framework 5 grant—GeneXtra.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.F. (e-mail: r.festenstein@imperial.ac.uk).

On the trail of SNPs

Extracting useful data from the human genome sequence is a major challenge. Lisa Melton examines the early steps towards personal genotyping.

GLAXOSMITHKLINE

One of the first approaches to extracting medical and biological value from the human genome has been the study and exploitation of single-nucleotide polymorphisms (SNPs; pronounced 'snips'), the most abundant type of genetic variation. SNP genotyping promises to reveal why some people are more susceptible to diseases such as cancer or diabetes, and what predisposes others to suffer adverse reactions to drugs. The implications for drug discovery, clinical drug trials, biomedical research, diagnostics and even forensics are considerable (see 'Crime busters', below, and 'Risk assessment', page 919).

To enable this promise to be fulfilled, faster and cheaper ways of screening individuals for SNPs are being developed. For start-up biotech firms and large pharmaceutical companies alike, the search for these genetic variations is now a serious business.

On average, two unrelated people differ at about 1 base in every 1,000 of the 3 billion or so bases in their genome, so any individual will have about 3 million SNPs.



Allen Roses hopes pharmacogenetics will help avoid adverse drug reactions.

What biomedical researchers are looking for are the much smaller subsets of SNPs that might be linked to the predisposition to disease or adverse drug responses. By comparing SNPs in patients and healthy controls it should be possible to track down those genetic differences. Once this

has been achieved, any person can be genotyped for this limited set of SNPs. Genotyping is a booming market: annual expenditure on SNP research is predicted to grow from US\$158 million in 2001 to more than \$1.2 billion in 2005.

As adverse drug reactions are the fifth-commonest cause of death in the United States, most major pharmaceutical companies have research and development (R&D) programmes to identify the SNPs that predispose to these responses. This pharmacogenetic strategy heralds an era in which the choice of drugs for a particular patient will be based on evidence rather than trial and error. "If you are going to sell a drug, what better characteristic to sell it on, than it will work for you and be safe," says Allen Roses, senior vice-president of genetics research at GlaxoSmithKline in Research Triangle Park, North Carolina. "By going along this pharmacogenetics route we will have more drugs out there that work and are safe to take," he says. To this end, around 1.8 million SNPs have been mapped with reference to the human genome sequence and placed in a public database by

CRIME BUSTERS

The ability to single out the Y chromosome and genotype male-specific markers is proving popular among forensic pathologists, because fully sequencing each DNA sample is prohibitively expensive. "With our Y genotyping kit, it is possible to screen many pieces of evidence at a relatively modest cost," says James Lazar, vice-president of research and development at Marligen Biosciences in Ijamsville, Maryland.

Marligen's Signet Y-SNP Identification System detects 43 single-

nucleotide polymorphisms

(SNPs) from the Y chromosome simultaneously. If these fail to match those of a suspect, he can be eliminated from consideration. And in rape cases, the ability to single out male DNA is a boon. Population genetics is another application.

"All SNPs in our panel have been widely studied and there is a lot of literature on the frequency of those SNPs in different populations. So one can figure out the ethnic background in a broad sense

by looking at those SNPs," says Lazar.

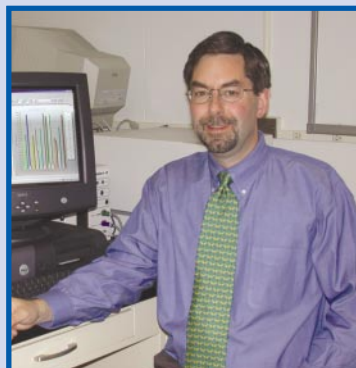
When samples have been out in the woods for weeks, or only a piece of bone is found, nuclear DNA can be degraded, yet the more plentiful mitochondrial DNA remains intact. A full forensic analysis of mitochondrial DNA can cost thousands of dollars, but Marligen provides an inexpensive alternative: the Signet Mitochondrial DNA Screening System, which analyses 30 variations at a cost of less than \$20 per sample.

Marligen's products all use the xMAP technology developed by Luminex in Austin, Texas. This combines microbead-based assays with detection by small lasers. The current platform uses 100 different colour-coded microspheres, to which capture probes are coupled. For many studies an array platform with 100 features is quite enough for large-scale high throughput. "It turns out you are usually only interested in screening a few dozen SNPs rather than tens of thousands," says Jim Jacobson, vice-president of research and development at Luminex.

The small number of features was decisive for Marligen. "In many disease states it is emerging that 20–50 genes are important. That's where we expect science to take us — more than one gene but less than 100," says Lazar.

The Luminex system allows multiplexing of up to 50 SNPs in a single well, saving up to tenfold on reagents and allowing the detection of 10,000 SNPs a day. "It's highly reproducible and very fast. To read a 96-well plate takes 30–40 minutes," says Lazar.

MARLIGEN



James Lazar: genotyping the Y.

L.M.

the not-for-profit SNP Consortium.

Of the more than 4 million SNPs mapped so far and deposited in public and private databases, only 3% are within genes. "People are not interested in all the polymorphisms; they want to look at those that are functionally relevant," says Dennis Gilbert, vice-president of genomics applications at Applied Biosystems in Foster City, California. This generally means those SNPs within genes and their regulatory regions. The company's Assays-by-Design service offers a route to this goal — it delivers pre-tested, ready-to-use assays for a customer's chosen SNPs.

Customers send in their target DNA sequence, with the SNP locations indicated. Applied Biosystems then develops optimal primers and 'TaqMan' probes to detect both alleles of the selected SNP sequences. Customers receive the probes for each SNP in a single tube with a two-dimensional bar code for sample tracking.

The assay is designed to run on Applied Biosystems' Sequence Detection Systems, real-time quantitative instruments for the polymerase chain reaction (PCR) that use the company's TaqMan probe technology. The use of a non-fluorescent quencher dye increases assay performance to the point at which it is possible to score 1,000 samples for 250 SNPs — that is, 250,000 genotypes — in a day, the company says.

Applied Biosystems also sells complementary off-the-shelf assays and

currently has more than 120,000 validated assays for SNPs in and around genes for linkage disequilibrium mapping studies. "You can go to our public website and type in a gene or region of the genome, and get a whole list of SNP assays for your association studies," says Gilbert.

The more the merrier

For those intending to genotype on a massive scale, Illumina of San Diego, California, has recently launched its hybridization-based BeadArray system. Each high-density array of probe-coated microbeads fits into a well of a standard 96-well microtitre plate and can simultaneously detect up to 1,152 different SNPs in a single DNA sample. The fully automated system can process more than a million assays a day. And reduced reagent volumes mean considerable savings in running costs, says Jay Flatley, Illumina's president and chief executive. "The strategy is to bring down the costs of high-throughput and highly accurate genotyping," he says.

The first system was sold to Génome Québec, a member of the International HapMap Project which is mapping SNP haplotypes in the human genome (see 'Haplotypes or pools?', page 921). The UK arm of HapMap, at the Wellcome Trust Sanger Institute in Hinxton near Cambridge, will also use Illumina's technology, and last autumn the US



APPLERA

Dennis Gilbert: online shopping for SNP assays.

National Institutes of Health awarded the company \$9 million to genotype 15% of the HapMap itself.

Illumina's million-dollar system is beyond the means of most academic researchers, but using the company's custom services, a screen for 2,000 SNPs, for example, would cost some 35 cents for each SNP assayed. "As the studies get larger the price drops considerably," says Flatley. "People who run our system can genotype at less than a nickel a data point."

The company has also developed a fine-mapping product — an extensive set of validated SNP assays — from a large genotyping study in collaboration with

RISK ASSESSMENT

THIRD WAVE TECHNOLOGIES

Even small numbers of single-nucleotide polymorphisms (SNPs) can work as predictive tools to assist clinical decisions, and this is the area that interests Third Wave Technologies. Based in Madison, Wisconsin, the company has set low-volume SNP-genotyping tests for individual diagnostics as its priority.

"We started out in the cardiovascular arena and helped to create the field for 'economy-class syndrome' mutation detection," explains chief executive Lance Fors. Third Wave developed a reasonably priced test for mutations in the blood-clotting factors V and II that can trigger venous blood clots in the brain and legs. Women who carry the factor V mutation (V Leiden) and are also taking contraceptive pills are at 100 times greater risk of deep-vein thrombosis and stroke, and today an increasing number of women taking contraceptives are tested for these variants. "It's moving towards mainstream medicine," says Fors. "An individual test for a specific SNP may turn into a \$10- or \$30-million market."

Low false negatives and low false positives are the winning solution in diagnostics, Fors adds. Third Wave's platform uses the firm's Invader RNA Assay. Two probes are designed to overlap or 'invade' each other to make a



Lance Fors: a future for predictive testing.

particular three-dimensional structure. If the mutation is present, there is an invasion and the structure is formed. A nuclease then recognizes this overlapping complex and cuts it. With 'normal' DNA, the probes cannot invade and the DNA remains intact, Fors explains. Structure-specific cleavage is far more precise than conventional methods based on the polymerase chain reaction (PCR), so mistakes are very rare, he points out.

The Invader technology is a one-step assay that simply requires the sample to be added. "Some of our clinical customers could never adopt PCR, because they couldn't afford the equipment or the infrastructure required to minimize contamination," Fors says.

Invader technology can scale up to detecting hundreds of thousands of SNPs, as its fluorescence readout is readily adaptable for high-throughput. It was used to screen 768 Japanese cardiovascular patients for the presence of 92,000 SNPs located in genes that are thought to confer a risk of cardiovascular disease, and to identify a candidate gene associated with susceptibility to heart attack on chromosome 6p21 (K. Ozaki *et al.* *Nature Genet.* **32**, 650–654; 2002).

L.M.

GlaxoSmithKline. These consist of about 42,000 SNP positions spread across the genome. Customers can buy them in localized pieces or they can run all 42,000 across their samples to get a broader association study.

The full monty

"The biggest problem is picking the right genes — it's a bit like playing the lottery," says Brad Margus, chief executive at Perlegen Sciences of Mountain View, California. "The other problem is that even if you find one gene associated with the disease, it may only account for 2% of the disease, so it can be very frustrating."

Whole-genome association studies that identify a set of SNPs linked with a disease are the obvious solution, but only a couple of years ago the cost would have been prohibitive. Today, Perlegen performs whole-genome studies at prices that few could have predicted. "Our people have devised a way to look at the entire genome at very high resolution," says Margus.

Formed in 2000 as an affiliate of DNA-microarray manufacturer Affymetrix in Santa Clara, California, Perlegen has taken the Affymetrix chip and shrunk the features to produce a glass wafer that holds about 60 million DNA probes — in contrast to the 500,000 of a standard Affymetrix array. In addition, Perlegen has developed more than 242,000 different PCR assays that amplify the entire genome.

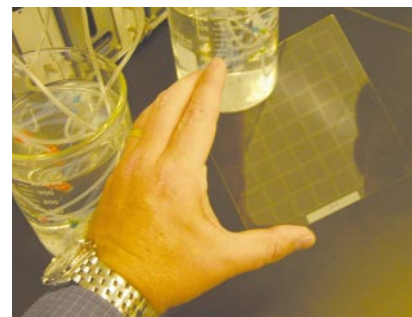
Compared with the few hundred bases usually amplified in conventional PCR reactions, the long-range assays average about 10,000 bases a time. These innovations, coupled with new high-resolution scanners, give a system that has allowed Perlegen to sequence the non-repeat regions of 50 haploid human genomes fully in just 15 months.

"Humans are very similar, so instead of looking at 3 billion bases in people who have a disease and those who don't, we sequenced 50 complete haploid genomes, which yielded 1.7 million SNPs," Margus explains. "We then designed SNP-reading oligonucleotide arrays to genotype those 1.7 million SNPs." With the Perlegen platform, the cost of genotyping these SNPs in 1,000 people is roughly \$2 million.

Lab in a tube

For academic researchers on a more modest budget, genotyping options are also opening up. ParAllele BioScience of South San Francisco, California, has developed a single-tube, amplification-based methodology, which it claims can detect up to 8,000 SNPs simultaneously. "You have one tube with 8,000 of these probes and you just add genomic DNA," says Tom Willis, ParAllele's chief executive. "The magic happens because we use a pool of molecular 'inversion probes' that target thousands of different SNPs in that one patient's DNA."

ParAllele's inversion probes are



Perlegen's chip holds 60 million probes.

insensitive to cross-reactions — often the limiting factor in highly multiplexed PCR reactions. The probes are incubated with the DNA (only 0.1 ng DNA per SNP is needed) and then all those that have bound are simultaneously targeted by a standard PCR reaction, using a common primer. "It's genotyping first and then amplifying all the SNPs at once," says Willis. The read-out signal is provided when the probe interacts with the genome and rearranges itself to create an amplifiable molecule. Once the assay is performed, the PCR products are loaded onto any chip platform for detection.

ParAllele claims that its technology achieves comprehensive coverage of the genome, or sections of the genome, at affordable prices, by reducing the millions of individual screening reactions typically

HAPLOTYPES OR POOLS?

As the Human Genome Project winds down, many researchers are stepping over into the next international public endeavour, the haplotype map project, or HapMap. This US\$100-million initiative was launched in October 2002 by the US National Human Genome Research Institute to provide the tools that will speed up the discovery of genetic factors that contribute to common conditions such as diabetes and heart disease. The results will be made quickly and freely available on the Internet.

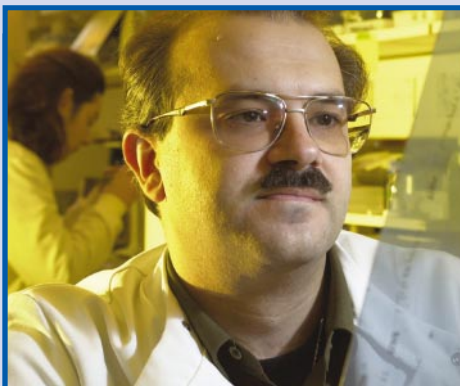
Along the genome, certain combinations of alleles at nearby SNP sites have remained unchanged over many generations and are inherited together as haplotypes. The aim of HapMap is to build a genome map of these haplotype blocks from different ethnic groups, mainly from China, Japan, Nigeria and northern Europe.

Labs in Canada, China, Japan, the United Kingdom and the United States will be involved. "The goal of the project is to validate SNPs across the entire genome and build tools to enable association studies, to compare genetic markers in people who have a disease with markers in a control population," says Panos Deloukas, senior group leader at the human genetics department at the Wellcome Trust Sanger Institute near

Cambridge, UK. The HapMap, set up as an international public-private research consortium, is expected to take three years to complete.

Why bother with a HapMap when whole-genome scanning technologies already exist? "With the technologies we have now, scanning the whole genome gene by gene is very costly. Scanning the genome haplotype by haplotype would be 5% of the cost," points out Julio Licinio at the pharmacogenomics lab of the University of California, Los Angeles.

An alternative way of reducing the costs of whole-genome scanning has just been launched by the discovery-genetics company Sequenom in San Diego, California. Its Allelotyping technology uses high-resolution MALDI-TOF (matrix-assisted laser desorption/ionization-time-of-flight) mass spectrometry to analyse the frequency of a SNP allele in large DNA pools of up to 500 individual samples. Only 50 picograms of DNA from each individual are required, and allelic frequencies as low as 3% can be detected. "Our customers can afford to do whole-genome scans today, because whole-genome scans are affordable on pools but not on individuals," says Charles Cantor, chief scientific officer at Sequenom.



Panos Deloukas: mapping SNP haplotypes.

used by PCR technology to a single reaction for each DNA sample. The system also boasts a high conversion rate — 85% of all SNPs are detected. “It allows researchers to pick which SNPs they want to interrogate, rather than having to use the SNPs that just happen to work,” says Willis.

The system has won the company a \$4-million grant from the US National Human Genome Research Institute to participate in the HapMap project in partnership with the Baylor College of Medicine in Houston, Texas.

Array-based genotyping

Affymetrix also aims to bring large-scale genotyping by DNA microarray within the reach of individual researchers. Currently in development is the GeneChip Human Mapping 10K Array, which should genotype more than 10,000 SNPs by allele-specific hybridization using only 250 ng of genomic DNA.

Key to this method is the preparation of the sample DNA. Using two standard molecular biology techniques, the genomic DNA is converted to a form that allows multiple loci to be amplified using a single PCR primer. The PCR product contains information for more than 10,000 SNPs which are then genotyped by hybridization to a single array.

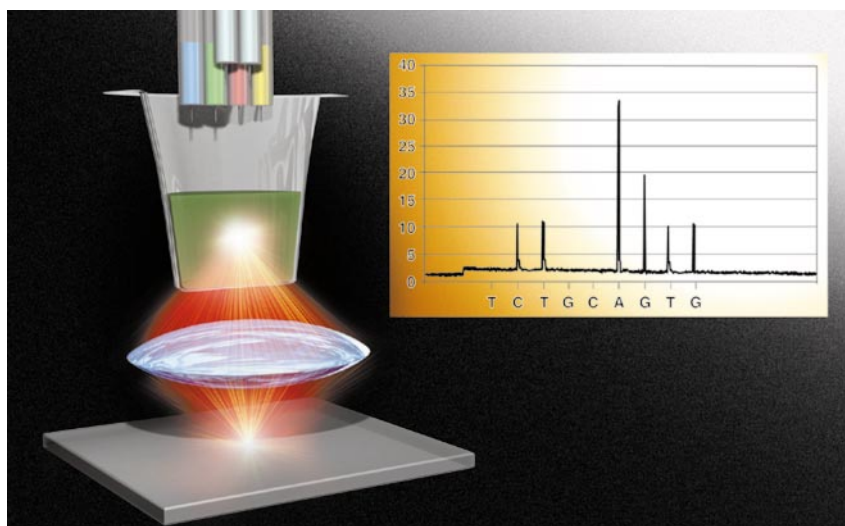
In addition to the SNPs the lab digests throw up, the draft human genome sequence makes it possible to run computer models of digests of total genomic DNA. “This allows us to predict which fragments we can amplify using this method and which SNPs reside in each fragment,” explains Keith Jones, vice-president for molecular genetics at Affymetrix.

The method can be used both for familial-based linkage analysis and for population genetic studies, for example. The assay can scale both downwards and upwards towards 100,000 SNPs, simply by changing the restriction enzyme used to prepare the sample.

One at a time

Sequencing is an important but often overlooked technology when it comes to scoring SNPs. “Most platforms fail to genotype 10–15% of the polymorphisms you are interested in. To get them typed, the best back-up methodology for us has been DNA sequencing at high reagent dilutions,” says Walt Klimecki, head of genomics at the Arizona Respiratory Center in Tucson.

The ability to sequence SNPs is what many customers of Pyrosequencing in Uppsala, Sweden, find attractive. “The advantage is you have the sequence context so you know you are looking at the right



Pyrosequencing throws light on sequence-based genotyping.

SNP,” says Karen Frost, the company’s market communications director. The technique is simple to use and delivers accurate and consistent analyses of a large number of genetic variations, she says. Because there is no restriction on where the primer can be positioned, it is possible to develop assays for all SNPs. “We have yet to find a SNP that we cannot design an assay for,” claims Frost.

The company’s technology is based on light emission. A sequencing primer is hybridized to the template DNA and the nucleotides are added one by one. If the nucleotide is complementary to the template, it is incorporated and light is emitted. If it is not complementary, there is no incorporation and no light.

Pyrosequencing’s workstations can analyse di-, tri- and tetra-allelic SNPs, insertions and deletions, as well as quantifying alleles in pooled samples. Other applications, such as bacterial typing, use the systems’ capacity for real-time, short-to-medium-length DNA sequencing. The company claims that the high-throughput automated version, with plate stacker and robotic arm, can score around 30,000 polymorphisms in 24 hours.

High-density maps

In contrast to the fluorescence-based readout systems used by most genotyping platforms, Sequenom, the discovery genetics company based in San Diego, California, uses MALDI-TOF (matrix-assisted laser desorption/ionization–time-of-flight) mass spectrometry in its SNP-discovery system and its Allelotyping whole-genome scanning system (see ‘Haplotypes or pools’, page 921).

“Fluorescence is a two-digit measurement, whereas spectroscopy is a four- or five-

digit measurement,” says Charles Cantor, Sequenom’s chief scientific officer. “We have very sharp peaks coming up against a relatively flat background, whereas in fluorescence the peaks are broad and tend to overlap.” The samples are presented to the mass spectrometer on a chip. “This lets us use very small samples — 100 times smaller than traditionally used — and also lets us automate,” Cantor explains.

The company has stockpiled 100,000 validated and immediately usable SNP assays through projects with the US National Cancer Institute and GlaxoSmithKline. “These are the gold standard set — the ‘try me first’ set,” says Cantor. “Most of them are near genes or in genes, the polymorphisms are real, the assays work and they are polymorphic enough to be cost-effective.”

With this constellation of ingenious approaches, SNP hunters on both the large and small scale are paving a fast track to personalized medicine. “The technologies have only just started to emerge and we are using them already — we do high-density whole-genome SNP mapping,” says Mihael Polymeropoulos, head of Novartis Pharma’s pharmacogenetics department in Gaithersburg, Maryland. “We now systematically collect DNA from every patient in every clinical trial, analyse that for variations and then at the end of the trial do association studies between the genetic variation, efficacy and the adverse effects.” Detecting SNPs is no longer the technical difficulty, the challenge is now to make sense of all the data. ■

Lisa Melton is science writer at the Novartis Foundation, London.

The SNP Consortium snp.cshl.org
HapMap hapmap.cshl.org

Company	Products/activity	Location	URL
Genotyping systems and services			
Aclara BioSciences	LabCard microfluidics arrays for multiplexed SNP genotyping	Mountain View, California	www.aclara.com
Affymetrix	GeneChip microarray systems, including the HuSNP mapping assay	Santa Clara, California	www.affymetrix.com
Amersham Biosciences	Instruments, equipment and products for genomics, proteomics and cellular assays; MegaBACE capillary-based, high-throughput automated sequencing and genotyping systems; CodeLink SNP bioarray for SNP genotyping and pharmacogenetic profiling	Uppsala, Sweden	www5.amershambiosciences.com
Applied Biosystems	DNA sequencers, mass spectrometers, and workstations and automated instruments for genomics and proteomics; Assays-by-Design and Assays-on-Demand genotyping services	Foster City, California	home.appliedbiosystems.com
Asper Biotech	Microarray and genotyping hardware and software	Tartu, Estonia	www.asperbio.com
BioChip Technologies	Pharm-O-Kin Screen; customized genotyping service based on biochip technology	Freiburg im Breisgau, Germany	www.genescan.com
deCODE Genetics	Pharmacogenomics services, high-throughput PCR-based microsatellite genotyping	Reykjavik, Iceland	www.decode.com
Exiqon	GenoView SNP genotyping assay	Vedbaek, Denmark	www.exiqon.com
Genome Therapeutics	GenomeVision Services for genomics including SNP discovery, validation and screening	Waltham, Massachusetts	www.genomecorp.com
Genometrix	Genotyping, gene expression and bioinformatics services	The Woodlands, Texas	www.genometrix.com
Illumina	Oligator custom DNA synthesis services; SNP genotyping services	San Diego, California	www.illumina.com
Ingenotyping	Custom transgenic mouse models	Martinsried, Germany	www.ingenotyping.com
Lambda	Biochips for genotyping and SNP detection	Freistadt, Austria	www.lambda.at
LGC	Pharmacogenetic diagnostics, consultancy and research services	Teddington, UK	www.lgc.co.uk
Marligen Biosciences	Off-the-shelf and customized SNP genotyping systems and services; DNA-purification products	Ijamsville, Maryland	www.marligen.com
MJ Research	Thermal cyclers, BaseStation DNA fragment analyser (sequencer and genotyper), labware, reagents for molecular biology	Waltham, Massachusetts	www.mjr.com
Myriad Genetics	Genomics-based drug discovery and predictive genetic testing	Salt Lake City, Utah	www.myriad.com
Nanogen	NanoChip automated workstation for SNP and STR analysis	San Diego, California	www.nanogen.com
Orchid BioSciences	Array-based SNP genotyping platforms: instruments and consumable kits for SNP scoring	Princeton, New Jersey	www.orchid.com
ParAllele Bioscience	SNP genotyping, genome variation scanning, gene-expression analysis	South San Francisco, California	www.p-gene.com
PerkinElmer Life Sciences	Automated systems for liquid handling and sample preparation; Micromax human DNA microarrays; equipment for PCR; AxyoPrime-FP SNP detection systems	Boston, Massachusetts	lifesciences.perkinelmer.com
Perlegen Sciences	Microarray-based SNP genotyping	Mountain View, California	www.perlegen.com
Professional Genetics Laboratory	Genotyping services for pharmacogenomics in clinical trials	Uppsala, Sweden	www.capio.se
Promega	READIT SNP genotyping system and associated products	Madison, Wisconsin	www.promega.com
Proteodyne	BioCube automated systems for SNP detection	Windsor, Connecticut	www.proteodyne.com
Pyrosequencing	PSQ 96MA and PSQ HS96A sequencing-based genotyping workstations for SNP screening, mutation detection and bacterial and viral typing	Uppsala, Sweden	www.pyrosequencing.com
QIAGEN	SNP analysis, molecular biology workstations	Valencia, California	www.qiagen.com
Sequenom	MassArray PCR-based system for high-throughput genotyping, targeted SNP discovery, allele-frequency analysis	San Diego, California	www.sequenom.com
Third Wave Technologies	Invader RNA assays and genotyping for a wide variety of human genes	Madison, Wisconsin	www.twt.com
Transgenomic	WAVE 3500 and WAVE-MD automated systems for mutation and polymorphism detection	Omaha, Nebraska	www.transgenomic.com
Variagenics	Products and services for genotyping and haplotyping for clinical trials	Cambridge, Massachusetts	www.variagenics.com

Instruments, systems and reagents for microarrays

Alpha Innotech	AlphaArray microarray reader	San Leandro, California	www.alphainnotech.com
Alpha Technology	Customized low- and medium-density microarrays	Hürth, Germany	www.alpha-tec.net
Axon Instruments	GenePix 4000B microarray scanner and GenePix Pro microarray analysis software	Foster City, California	www.axon.com
Coming	GAPS II microarray slides, hybridization chambers	Coming, New York	www.coming.com
febit	Geniom one integrated platform for DNA microarrays	Mannheim, Germany	www.febit.com
GeneScan Europe	High through-put production of low- and medium-density microarrays	Freiburg, Germany	www.genescan-europe.com
Genomic Solutions	GeneTAC biochip system	Ann Arbor, Michigan	www.genomicsolutions.com
Mosaic Technologies	EZ-RAYS activated slide kits for DNA microarrays	Waltham, Massachusetts	www.mostek.com
Packard Bioscience	Hydrogel-coated slides, arrayers, scanner and data-analysis software	Meriden, Connecticut	www.packardbioscience.com
PicoRapid Technologie	PicoArrays, PicoSlides and microarray spotting service	Bremen, Germany	www.picorapid.de
Scandinavian Micro Biodevices	Slides for DNA arrays	Lynby, Denmark	www.smb.dk
Scienion	Generic and customized microarrays	Berlin, Germany	www.scienion.de
Sigma-Genosys	Custom oligonucleotides, oligonucleotide libraries, DNA arrays	The Woodlands, Texas	www.sigma-genosys.com
SurModics	3D-Link activated slides	Eden Prairie, Minnesota	www.surmodics.com
TeleChem/ArrayIt International	Distributor of fine chemicals and manufacturer of ArrayIt products for microarray analysis	Sunnyvale, California	www.arrayit.com
VBC-GENOMICS	DNA sequencing, custom oligonucleotide microarrays and Affymetrix GeneChip services	Wien, Austria	www.vbc-genomics.com
Vysis	GenoSensor microarray readers and reagents	Downers Grove, Illinois	www.vysis.com
Zeptosens	SensiChip microarray systems	Witterswil, Switzerland	www.zeptosens.com

Instruments, systems and reagents for PCR

BD Biosciences	Reagents for PCR	Franklin Lakes, New Jersey	www.clontech.com/index.shtml
Bio-Rad	Products, instruments and software for life-sciences research	Hercules, California	www.bio-rad.com
Cepheid	Instruments and equipment for DNA purification and PCR	Sunnyvale, California	www.cepheid.com
Corbett Research	Equipment for DNA analysis and PCR	Sydney, Australia	www.corbettresearch.com
Epoch Biosciences	Custom oligonucleotides for PCR and microarrays; MGB Eclipse probe systems for gene-expression quantitation, SNP detection and pharmacogenomic studies	Bothell, Washington	www.epochbio.com

Company	Products/activity	Location	URL
Genes 4-U	Validated ready-to-use ToolSets for mutation detection using real-time fluorescent PCR with the Roche LightCycler system	Zurich, Switzerland	www.genes-4u.com
Hamilton Robotics	PCR set-up, ultra-filtration and clean-up, plasmid purification and gDNA extraction	Bonaduz, Switzerland	www.hamiltoncompany.com/robotics
Idaho Technology	Thermocyclers, probes and reagents for PCR	Salt Lake City, Utah	www.idahotech.com
RNA-TEC	Custom oligonucleotide synthesis	Leuven, Belgium	www.rna-tec.com
Stratagene	Tools and reagents for genomics, proteomics, drug discovery and toxicology	La Jolla, California	www.stratagene.com
Takara Bio	PCR-related products, DNA microarrays, DNA sequencing service	Shiga, Japan	www.takara-bio.co.jp
Techne	Laboratory equipment; thermal cyclers	Princeton, New Jersey	www.techne.com
Thermo Hybaid	Instruments and equipment for molecular biology; custom peptide and oligonucleotide synthesis	Ashford, Middlesex	www.thermohyбайд.com
Thermolyne	Hot plates, mixers and stirrers	Dubuque, Iowa	www.barnsteadintl.com
TIB MOLBIOL	LightCycler hybridization probes, custom probe design, custom assay development	Berlin, Germany	www.tib-molbiol.com

DNA sequencing

GATC Biotech	High-throughput DNA sequencing and analysis, SAGE	Konstanz, Germany	www.gatc-biotech.com
Lark	DNA sequencing, gene detection, and genotyping services	Houston, Texas	www.lark.com

Software

Accelrys	GCG Wisconsin package for sequence and genome analysis; MacVector for sequence analysis; Discovery Studio products for database mining, genomics and proteomics	San Diego, California	www.accelrys.com
BioDiscovery	Software products for microarray research and analysis (CloneTracker, ImaGene, GeneSight, GeneDirector)	Los Angeles, California	www.biodiscovery.com
Biocomputing Platforms	Software for management of clinical genotyping data	Espoo, Finland	www.biocomputing.fi
Bruker Daltonics	GenoTools for high-throughput multiple biallelic SNP genotyping by MALDI-TOF mass spectrometry	Billerica, Massachusetts	www.bdal.com
Celera	Discovery System web-based informatics tools for accessing the Celera annotated genome databases; bioinformatics services; DNA sequencing	Rockville, Maryland	www.celera.com
LION Bioscience	Software for sequence and genome analysis; SRS-based DiscoveryCenter platform for integration of chemical and biological information; contract bioinformatics services	Heidelberg, Germany	www.lionbioscience.com
Nonlinear Dynamics	Image-analysis and database software for 1D and 2D gel electrophoresis and high-throughput arrays; analysis of high-throughput 2D gels for proteomics	Newcastle-upon-Tyne, UK	www.nonlinear.com
Paracel	AGENT automated genotyping tool for sequence-based analysis	Pasadena, California	www.paracel.com
Premier Biosoft	Desktop bioinformatics packages for sequence analysis, oligonucleotide and primer design, and two-hybrid protein interactions	Palo Alto, California	www.premierbiosoft.com
Progeny	Software for drawing pedigrees and compiling databases	South Bend, Indiana	www.progeny2000.com
Scanalytics	MicroArray Suite for microarray analysis; Gene Profiler gel-analysis software for genotyping	Fairfax, Virginia	www.scanalytics.com
Strand Genomics	Oligonucleotide design and microarray analysis software	Bangalore, India	mail.strandgenomics.com
SunErgia Group	Genorama genotyping software for analysis of microarray-based resequencing or genotyping images	Reston, Virginia	www.sunergiamedical.com
Visualize	Visual Genetics software for patient pedigrees, genotypes and clinical data	Phoenix, Arizona	www.visualizeinc.com

General

Agilent	Workstation for RNA, DNA or protein analysis; instruments for genomics and proteomics	Palo Alto, California	www.agilent.com
Apogent Technologies	Labware and equipment for the life sciences	Portsmouth, New Hampshire	www.apogent.com
BioCat	Kits and reagents for PCR and microarrays	Heidelberg, Germany	www.biocat.de
BMG Labtechnologies	FLUOstar OPTIMA and POLARstar OPTIMA plate readers	Offenburg, Germany	www.bmglabtech.com
Brinkmann	Laboratory instrument suppliers	Westbury, New York	www.brinkmann.com
Eppendorf	Laboratory instrumentation and consumables for molecular and cell biology	Hamburg, Germany	www.eppendorf.com
GeneMachines	Automated instruments for sample preparation, microarraying, DNA purification, oligonucleotide synthesis and colony/plaque picking	San Carlos, California	www.genemachines.com
Genetix	Picking, arraying, gridding, replicating, rearranging and microarraying robots for genomics and proteomics	New Milton, UK	www.genetix.co.uk
ICN Biomedicals	Reagents and biochemicals for life-sciences research	Aurora, Ohio	www.icnbiomed.com
Incyte Genomics	LifeSeq annotated gene and expressed sequence-tag databases; Proteome BioKnowledge Library species-specific protein information databases; bioinformatics software	Palo Alto, California	www.incyte.com
Invitrogen	Kits and reagents for genomics, proteomics, molecular and cell biology	Carlsbad, California	www.invitrogen.com
Luminex Corporation	xMAP platform for microbead-based multiplex assays; Luminex100 flow cytometer	Austin, Texas	www.luminexcorp.com
Millipore	Automated equipment for molecular biology, biochemistry, genomics and proteomics	Bedford, Massachusetts	www.millipore.com
MiraiBio	Luminex bead-based detection systems; FMBIO imaging system	Alameda, California	www.mirai.bio.com
MWG Biotech	Thermal cyclers, and automated systems for high-throughput sequencing and production of microarrays	Ebersberg, Germany	www.mwg-biotech.com
New England Biolabs	Reagents and kits for molecular biology	Beverly, Massachusetts	www.neb.com
Pierce Biotechnology	Components and consumables for automated systems in molecular biology, protein chemistry, sample handling, immunology and chromatography	Rockford, Illinois	www.piercenet.com
Roche Diagnostics	Reagents and kits for molecular biology, functional genomics and proteomics research	Lewes, UK	www.roche-applied-science.com
Sigma-Aldrich	Reagents and biochemicals for life-sciences research	St Louis, Missouri	www.sigmaaldrich.com
USB	Biochemicals and reagents for molecular biology, including PCR, manual DNA sequencing, mutation detection	Cleveland, Ohio	www.usbweb.com
Westburg	Products for PCR, microarrays and molecular biology	Leusden, The Netherlands	www.westburg.nl

● See advertisement

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building

4 Crinan Street

London N1 9XW, UK

Tel +44 (0) 20 7843 4961

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Silje Opstrup (4994)

France/ Belgium:

Amelie Pequignot (4974)

Production Manager: Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Switzerland:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikatamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Hideki Watanabe

e-mail: h.watanabe@naturejp.com

naturejobs

Beyond the helix

The 50 years that have passed since DNA's double-helix structure was elucidated could be commemorated with skills that have evolved from progressive to *passé*. Pipetting? Prosaic. Gene cloning? Available in off-the-shelf kits. Whole-genome sequencing? Reserved for a handful of high-throughput, highly automated facilities. So what are the new skills or techniques that will keep molecular biologists in funding for the next few years?

Eric Green, director of the US Intramural Sequencing Center in Bethesda, Maryland, sees bioinformatics as the next frontier, but he quickly amends himself, saying that the conception of what bioinformatics was and what it is becoming is changing rapidly. Originally it involved the skills needed to find single genes in databases. The future will require scientists to manipulate large data sets, often moving from one to another and drawing parallels between them.

Eric Lander, director of the Whitehead Institute's Center for Genome Research in Cambridge, Massachusetts, forecasts a specific application for this approach. With several mammalian genomes soon to be completed, the future holds the prospect of cross-comparison of the various sequences, which will help researchers to identify what functions genes perform. There will be work for people who create the informatics tools, Lander says, as well as for the people who use them.

Perhaps the best way for molecular biologists to further their careers is to remember their history, but keep their perspective. Green notes that tools are a means and not an end. In the 1970s and 1980s, he says, lab heads didn't advertise for a pipetter, they searched for molecular biologists. So too, in the future, scientific directors will not seek bioinformaticians, but scientists who use the latest databases, software and algorithms as tools.

Paul Smaglik

Naturejobs editor



Contents

SPECIAL REPORT

Shedding light on
stem-cell research

p928

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

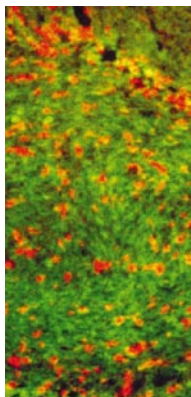
FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS



Complex political, ethical and legal issues surround research on human embryonic stem cells. Diane Gershon explores the field's long-term career prospects.

Human embryonic stem (ES) cells can give rise to specialized cell types of any kind, such as neurons or insulin-producing β cells. As such, many claims have been made about the potential benefits of research involving human ES cells and their use to treat conditions such as Parkinson's disease, spinal-cord injury and diabetes. In reality, researchers are only just beginning to scratch the surface. But legal, political and ethical questions surround work on stem cells derived from human embryos, leaving researchers with an uncertain future.

Many observers predicted that there would be a wholesale exodus of US scientific talent following the decision by President George W. Bush to ban the use of federal funds for research on human ES-cell lines

derived after August 2001. But with states such as California enacting legislation in support of the research and rolling out the red carpet to researchers, investors and companies alike, such an exodus has not occurred — at least not for now.

Legislation governing human ES-cell research varies widely from country to country, even within the European Union (EU). If an EU vote on 10 April to restrict research involving human ES cells was to become law, then

this avenue of research might be closed down — even in countries such as the United Kingdom, which currently has one of the more liberal and favourable research environments for the field. It is joined in this position by Australia, China, India, Israel, Singapore and Sweden.

The government in Singapore, for example, sees human ES-cell research as having considerable commercial potential, and is actively promoting and funding work in this area. It matches well with a recent mandate to train more Singaporeans in the life sciences, says Ariff Bongso, an *in vitro* fertilization (IVF) specialist at the country's National University Hospital. "Human ES-cell research in Singapore is progressive, enthusiastic and encouraging," he says.

Bongso, who trained in North America, notes that he has recently been on the receiving end of increased interest from US scientists who are seeking collaborations outside the United States to gain access

to cell lines and technical know-how.

Austin Smith, head of the Institute for Stem Cell Research at the University of Edinburgh, UK, says that he too has received more approaches over the past year from postdocs in other countries. "They feel that things are more open here," he says.

That may be true, but there is still UK bureaucracy to cope with in order to secure a licence from the Human Fertilisation and Embryology Authority to derive new human ES-cell lines for research, says Stephen Minger of the GKT Centre for Neuroscience Research at King's College London. Although it is legal in Britain to create human embryos for stem-cell research using nuclear transplantation, the three projects that have been granted licences so far all use surplus embryos from IVF clinics.

Securing a licence can be an onerous and lengthy procedure, and you have to be willing to put yourself under regulatory and administrative scrutiny, says Minger. But it is the right way to go about it, he says. "The policies here are very clear-cut; it doesn't hinder us in any way," he explains. Minger came to Britain from the United States as a postdoc about seven years ago and now runs a neural stem-cell group that is increasingly moving into human ES-cell research.

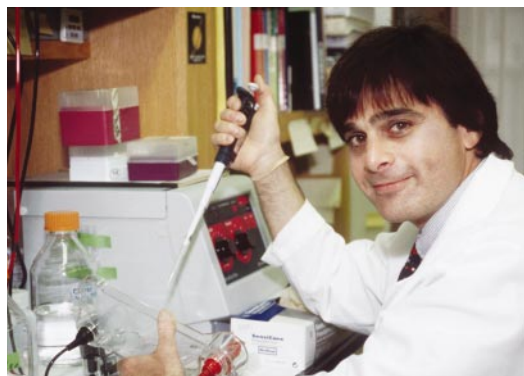
Lack of a steady supply of embryos is another issue, says Minger. Moreover, he says, the work itself is technically demanding and labour-intensive. The field is still in its infancy and no one has really mastered it, so there is little technical experience to draw upon (see 'Education, education, education', opposite).

THROUGH THE LOOPHOLE

Despite the restrictions within the United States — and the looming threat of an outright ban on all forms of human cloning, including for 'therapeutic' use — there are regions that offer a favourable climate for stem-cell research. California, for example, has passed legislation in support of human ES-cell research in the hope that this will allow it to recruit and retain high-flying scientists, as well as helping to attract investors and biotechnology companies.

Some US institutions are also exploiting a loophole in the current Bush policy. Although new human ES-cell lines cannot be derived using federal funding, the policy does not preclude the use of private or state funds. Several foundations have already stepped up

Evan Snyder is certain that California will attract talented researchers for stem-cell work.



their involvement. Last June, for instance, the Juvenile Diabetes Research Foundation International in New York launched a US\$20-million fundraising effort in support of human ES-cell research worldwide.

Many Californian institutions are now seeking creative ways to fund human ES-cell research from non-federal sources, in particular by setting up independent programmes. Evan Snyder, who recently left Harvard Medical School to join the Burnham Institute in La Jolla as director of its new stem cell and regeneration programme, believes that the conditions in California will attract talented individuals. Mark Mercola, who studies cardiac development in *Xenopus* embryos and is also ex-Harvard, has already joined him there. Snyder says that he is seeking researchers with a firm grasp of developmental biology, particularly cellular and molecular developmental biology of any system, to staff the new programme.

But he may have competition. Last August, the University of California, San Francisco, announced plans to establish a Developmental and Stem Cell Biology Program using a \$5-million grant from Andy Grove, the chairman of Intel. In December, Stanford University followed suit, revealing plans for an Institute for Cancer/Stem Cell Biology and Medicine to be set up with \$12 million in seed money from an anonymous donor. Other states, including Massachusetts and New Jersey, are now considering legislation similar to that in California to authorize stem-cell research.

In contrast to the United States, the federal parliament in Australia voted last year to allow ES-cell research using excess IVF embryos. The government is providing A\$43.5 million (US\$26.3 million) in federal funding to establish a Centre for Stem Cells and Tissue Repair at Monash University in Victoria. The centre will consolidate efforts in stem-cell research across Australia and will be headed by Alan Trounson.

With a staff of nearly 140 when it opens later this year — two-thirds of whom will be PhDs — Trounson says that the centre will offer great opportunities for established and young scientists interested in adult and ES-cell development and differentiation, tissue engineering and transplantation biology. Already it is proving attractive — Stephen Livesey is returning to Australia after a 14-year stint in the United States to become the centre's director of tissue regeneration.

For researchers elsewhere, such opportunities must seem slightly galling. Oliver Brüstle, a neuroscientist at

Education, education, education

To train more scientists in the lab techniques that are used in human embryonic stem (ES)-cell research, the US and UK governments are setting up a number of training courses and career-development opportunities.

The US National Institutes of Health plans to fund five institutions to develop training opportunities, including short courses to provide hands-on training in human ES-cell culture techniques, directed cell differentiation, gene transfection and cell characterization.

The Jackson Laboratory in Bar

Harbor, Maine, for example, already runs an annual summer workshop on protocols in stem-cell biology. And starting this month, the WiCell Research Institute, a non-profit subsidiary of the Wisconsin Alumni Research Foundation, is offering two-day stem-cell training courses to be held monthly in Madison.

Britain's Medical Research Council also recognizes the need to train more scientists in this area. It recently awarded 11 studentships specifically in stem-cell research to individuals looking to pursue a career in this field.

D.G.

the University of Bonn in Germany, works under much tougher conditions. German law prohibits the use of funds — public or private — to derive new human ES-cell lines, and researchers can only work on cell lines created before January 2002. Even so, Brüstle says it took him more than two years to get permission to import human ES-cell lines from Israel as part of a collaboration with the Rambam Medical Center in Haifa.

Brüstle worries that such restrictions will prevent German researchers from participating fully in international programmes or taking advantage of new technological developments. They could also make it harder to stimulate biotech interest in this area, leaving Germany at an economic disadvantage.

Admitting that he seriously considered leaving his country for Britain 18 months ago, Brüstle says that he stayed partly because his grant application to use human ES cells made him something of a “test case”. Brüstle believes the ball is now in the scientists’ court to demonstrate the therapeutic potential of these unique cells. “The more data are accumulated, the easier it will be to convince the politicians,” he says, which will lead to fewer restrictions on who can do what and where. ■

Diane Gershon is Assistant Editor, New Technology, for Nature Medicine.

Austin Smith says that over the past year he has received more approaches for work than ever before from postdocs in other countries.

